**A Dissertation**

**For the Degree of Doctor of Philosophy**

# Studies on Developmental Genetic Programs of Early Embryogenesis in Chicken

# 닭의 초기 배아 발생의 발달 유전적 프로그램에 대한 연구

**February, 2018**

**By**

**Young Sun Hwang**

**Biomodulation Major**

**Department of Agricultural Biotechnology**

**Graduate School, Seoul National University**

# SUMMARY

Avian species have been utilized as valuable model system for several studies, especially for embryology. Due to their important position among the vertebrates in the comparative genomics of vertebrate species and wide relevance across many research fields, Bird10K project (B10K) was initiated during 2014~2015. As a part of this project, a phylogenetic hierarchy of avian species and the comparative genomics for flight and functional adaptations were generated. Although historically significant and theoretically excellent, there is no genome-wide transcriptomic data investigating the initial and critical events of avian embryogenesis, mainly because of practical difficulties in accessing the pre-oviposited embryos. In this regard, we presented the first whole transcriptome data of pre-oviposited embryos, including oocyte, zygote, and intrauterine embryos from Eyal-giladi and Kochav stage I (EGK.I) to EGK.X using a non-invasive method in chicken. Furthermore, our multi-omics approaches investigated the first transcriptional activation upon fertilization in chicken. Also, we identified avian-specific small heat shock protein, HSP25, and found its putative roles in the suspended early development, avian blastoderm dormancy.

The first study was performed to generate whole-transcriptomic information of early chicken embryos and analysis. In this study, a total of 137 pre-oviposited embryos from ovary and oviduct were collected, and RNA sequencing (RNA-seq) was performed in chicken. From these results, two waves of chicken zygotic genome activation (ZGA) governed the distinct developmental programs, such as Notch, MAPK, Wnt, and TGF-beta signaling, between cleavage and area pellucida formation period,

separately. Furthermore, EGK stages of chicken were mapped into counterparts in human and mouse, and highlighted the chicken-specific features of signaling pathways, as well as the gradually analogous expression pattern after ZGA. These findings provide a genome-wide understanding of avian embryogenesis and comparative studies among amniotes.

Based on the generated whole-transcriptome sequencing, we identified maternal-to-zygotic transition (MZT), which is a critical process for establishing embryonic identity throughout the vertebrates. Based on the co-expression analysis on the RNA-seq data according to semi-supervised clustering, two waves of ZGA-mediated MZT were observed across the intrauterine stages and transcriptional and translational dynamics were associated. Furthermore, the transitions were found according to the distinct developmental characteristics between cleavage and area pellucida formation period functionally. Finally, epigenetic reprogramming and miR-302s expression suggest that the definite MZT in EGK.VIII during early chicken development. Our study is expected to provide an evolutionary link among vertebrates in perspective of MZT regulation.

Induced pluripotent stem cells (iPSCs) and extended pluripotent stem cells (EPSCs) were generated in mammals, based on studies of the factors with transcriptional regulation in oocytes and early embryos. However, using well-known mammalian transcription factors (TFs) without examining the factors during early development in birds, the iPSC-like cells in avian species have not been fully reprogrammed. In the third study, we characterized the transcriptional transition and related TFs during early chicken development based on

a transcriptomic analysis. As a result, two waves of transcriptional activation were found and the TFs were provided in each wave prior to cleavage and area pellucida formation during intrauterine development. Our results contribute to identifying a number of reprogrammable TFs and demonstrate that the cleavage stage, showing a similar state with totipotency compared to mammals and higher-potent features than area pellucida formation period, is the pursuit of reprogramming in birds.

For looking more deeply into ZGA in avian species, we investigated the detailed insight into the onset of genome activation in chicken. Our transcriptomic analysis of early chicken embryos clarified two waves of transcriptional activation after fertilisation and Eyal–Giladi and Kochav Stage V (EGK.V) at both the genome-wide and base resolution of intronic regions. Furthermore, we determined the allelic expression based on breed-specific single nucleotide polymorphisms (SNPs) and traced the first-wave transcripts, accomplished via a multi-omics approach using genome resequencing of parents and whole-transcriptome analysis of their individual embryos. Surprisingly, maternal genome activation (MGA) was exclusively found in the zygote stage, regardless of the presence of haploid genomes in male PN and supernumerary sperm in the egg; this maternal gene expression regulated the cleavage period and was replaced by bi-allelic expression after MZT. Our results demonstrate that the feasible mechanism in avian zygotes activating only maternal alleles to inhibit disproportionate genome contribution or genetic instability due to inattentive transcription of paternal alleles.

Finally, we discovered the avian-specific small heat shock proteins in stress-tolerant blastoderm. We found that chicken *HSP25*

was expressed especially in blastoderm and highly upregulated during low temperature storage. Multiple alignment, phylogenetic trees and the expression in blastoderms of Japanese quail and zebra finch showed homologs of *HSP25* were conserved in avian species. After knockdown of chicken *HSP25* using siRNAs in blastodermal cells, pluripotency marker genes significantly decreased. Furthermore, this loss of function studies have demonstrated that chicken *HSP25* is associated with anti-apoptotic, anti-oxidant and pro-autophagic functions in chicken blastodermal cells. Collectively, avian *HSP25* is important in association with the very first line of cellular defense against environmental stresses, and protection of future embryonic cells in avian blastoderm.

Based on the researches, we provide the extensive early development such as ZGA, MZT, signaling pathways in an avian manner from the first transcriptomic approach. Furthermore, unlike the relatively conserved second-wave ZGA, the first-wave ZGA is showing the variable characteristics in each species, which may be due to evolutionary outcomes. Also, the avian-specific HSP25 is the constitutive protector in avian blastoderm dormancy. Our results will filling the gap for comparative studies among the species in developmental and evolutionary respects as a valuable resources for future availability, as well as contribute to comprehensive understanding early avian embryogenesis.

**Keywords:** aves, chicken, early embryo, next-generation sequencing, HSP25

**Student Number:** 2010-22862

# CONTENTS

# LIST OF FIGURES

## CHAPTER 3



## CHAPTER 4

**CHAPTER 5**



**CHAPTER 6**

**CHAPTER 7**

# LIST OF TABLES

## CHAPTER 3



## CHAPTER 4

**CHAPTER 5**

**CHAPTER 6**

**CHAPTER 7**

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **EGK** | Eyal-Giladi and Kochav |
| **HSPs** | Heat shock proteins |
| **RNA-seq** | RNA sequencing |
| **HH** | Hamburger and Hamilton |
| **ZGA** | Zygotic genome activation |
| **MZT** | Maternal-to-zygotic transition |
| **iPSCs** | Induced pluripotent stem cells |
| **EPSCs** | Extended pluripotent stem cells |
| **TFs** | Transcription factors |
| **SNPs** | Single nucleotide polymorphisms |
| **MGA** | Maternal genome activation |
| **NGS** | Next-generation sequencing |
| **PN** | Pronucleus |
| **BrdU** | Bromo-uridine |
| **miRNAs** | MicroRNAs |
| **rRNA** | Ribosomal RNA |
| **MBT** | Mid-blastula transition |
| **WL** | White leghorn |
| **KO** | Korean oge |
| **WTS** | Whole-transcriptome sequencing |
| **DAVID** | The Database for Annotation, Visualization and Integrated Discovery |

| | |
|---|---|
| **GO** | Gene ontology |
| **KEGG** | Kyoto Encyclopedia of Genes and Genomes |
| **GEO** | Gene Expression Omnibus |
| **MDS** | Multidimensional scaling |
| **FDR** | False discovery rate |
| **TMM** | Trimmed mean of M-values |
| **RPKM** | Reads per kilobase million |
| **DEGs** | Differentially expressed genes |
| **hpf** | Hours post-fertilization |
| **WCGNA** | Weighted correlation network analysis |
| **snoRNA** | Small nucleolar RNA |
| **p-PolII** | Phosphorylation of RNA polymerase II C-terminal domain |
| **2C** | 2-cell embryo |
| **Animal-TFDB** | Animal Transcription Factor Database |
| **PGCs** | Primordial germ cells |
| **WGS** | Whole-genome sequencing |
| **ALDB** | Domestic Animal Long Noncoding RNA Database |
| **pre-miR** | Pre-matured miRNA |
| **snRNA** | Small nuclear RNA |
| **lincRNA** | Long intergenic non-coding RNA |
| **SMART** | Simple Modular Architecture Research Tool |

# CHAPTER 1

# GENERAL INTRODUCTION

Avian species have been used as valuable model systems for virology, cancer and immunology (Stern, 2005). In particular, because the avian embryo develops in eggs before hatching, it is an excellent in vitro-like in vivo system; this has allowed much research on the developmental events during embryogenesis (Hamburger and Hamilton, 1951). Molecular genetic studies of primitive streak formation and gastrulation after oviposition have been performed widely in avian species, which have made a great contribution to embryology (Kispert et al., 1995, Streit et al., 2000, Skromne and Stern, 2001, Torlopp et al., 2014). Although theoretically excellent, practical difficulties in accessing the pre-oviposited eggs have resulted in only a few studies solely from a morphological perspective during early development after the criteria of Eyal-Giladi and Kochav (EGK) in 1976 (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980, Park et al., 2006). Recently, the limited molecular studies early embryogenesis in avian species have been evaluated using non-invasive methods acquiring intrauterine chick embryos, and zebra finch embryos at stages before EGK.X (Lee et al., 2013, Mak et al., 2015, Nagai et al., 2015).

With the development of next-generation sequencing technologies, the transcriptomic analysis in biological investigation have been established as a basic approach. Also in embryologic studies, investigating the temporal regulation of gene expression at transcriptome level is important for understanding the extensive early embryonic development. Transcriptional regulation of early developmental processes, including maternal mRNA degradation, induction of transcripts by ZGA, maternal to zygotic transition (MZT), cell signaling regulation, transcription factors, and epigenetic reprogramming, has already been characterized in representative model organisms, such as mice, cows, frogs and zebrafish, as well as humans, using next-generation sequencing technology (Aanes et al., 2011, Tan et al., 2013, Xue

et al., 2013, Yan et al., 2013, Cantone and Fisher, 2013, Graf et al., 2014, Lee et al., 2014). Despite their great value as models in developmental biology and their important position in evolution, no genome-wide study of developmental characteristics during the first 25 h of early avian development has been reported previously.

Through the intrauterine early development and following oviposition in avian species, avian embryos will experience the blastoderm dormancy, also called cold torpor. Either in nature or in poultry industry (Winkler and Walters, 1983, Gomez-de-Travecedo et al., 2014), the blastoderm dormancy is crucial point to be studied as an extended and another program in early avian embryogenesis. Particularly, species-specific small heat shock proteins (sHSPs) in function as the critical role in dormant embryos (Hsu et al., 2003, Ludewig et al., 2004, McElwee et al., 2004, Villeneuve et al., 2006, Qiu and Macrae, 2008, King and MacRae, 2012, MacRae, 2016). Although the recent study of extensive investigation about avian blastoderm dormancy was conducted, the avian-specific sHSP has not been studied yet.

To extensive understand the temporal expression during early development from fertilization to blastoderm dormancy in avian species, a series of analysis and experiement were conducted. In CHAPTER 2, we review the early developmental characteristics of avian species and other species. In addition, the suspended developmental programs in various species will also be discussed. In CHAPTER 3, we obtained early chicken embryos, generated whole-transcriptome RNA sequencing data, and signaling pathways for early avian development. In CHAPTER 4, we analyzed our transcriptome data differently to characterize maternal-to-zygotic transition in chicken. In CHAPTER 5, we parsed the genome-wide transcription factors and provide their

expression during early development in chicken. In CHAPTER 6, the multi-omics approaches were conducted to investigate avian-specific first-wave zygotic genome activation. Finally in CHAPTER 7, the avian-specific small heat shock protein for blastoderm dormancy is found and discussed.

# CHAPTER 2

# LITERATURE REVIEW

## 1. Early embryogenesis in chicken

### 1.1. The first day of embryonic development in chicken

After oviposition, chick embryo develops in eggs and needs approximately 21 days for hatching, *ex utero* (Hamburger and Hamilton, 1951). At the time egg is laid, chicken blastoderm contains 40,000-60,000 cells with a large amount of yolk (Eyal-Giladi and Kochav, 1976). Thus, the initial events of embryogenesis in chicken, such as fertilization, cell division, and lineage determination, occur prior to oviposition.

After ovulation, a matured oocyte is fertilized at the infundibulum. The first cleavage begins about 5 hours after fertilization (Olsen, 1942), and another development taking 20 hours takes place in the uterus, which is called intrauterine development. According to the criteria of Eyal-Giladi and Kochav (EGK) (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980), intrauterine development before oviposition in avian species includes ten stages, consisting of cleavage (EGK.I to EGK.VI) and the area pellucida formation period (EGK.VII to EGK.X). These two periods are clearly morphologically distinct: the former is characterized by rapid asymmetric cellularization, and the latter by axis formation and lineage segregation.

## 2. Fertilization and initiation of chicken development

### 2.1. Polyspermy fertilization in chicken

Compared to monospermic mammals, which suffer from embryonic lethality in pathological polyspermy due to the mitotic dysfunctions from an excessive sperm asters (microtubules), polyspermy mechanism is essential for

initiating fertilization in avian species (Snook et al., 2011). Chick embryos show normal development after the penetration of numerous sperms, suggesting that polyspermy may be involved in the first development in chicken (Birkhead et al., 1994). In polyspermy animals including avian and amphibian species, the long-lasting small $Ca^{2+}$ rise by multiple sperm finally result in inducing egg activation (Iwao, 2012). The polyspermy mechanism of fertilization seem to fulfill the need for a number of sperm to activate large eggs (Iwao, 2012). In addition, intracytoplasmic sperm injection (ICSI) with sperm extracts or their components including phospholipase Cζ (PLCZ), aconitate hydratase (AH) and citrate synthase (CS) generated the full developed quail, indicating the importance of polyspermy in fertilization in avian species (Mizushima et al., 2014).

In the intrauterine embryos, sperm dynamics in chicken were examined in the previous study (Lee et al., 2013a). Three types of nucleus were found in zygotic formation, condensed, and decondensed supernumerary sperms. Regardless of paternal pronuclear, condensed sperm heads were mainly located in the perivitelline space and cytoplasm, but hardly in the yolk region. Whereas decondensed sperm heads were found in the yolk only. These supernumerary sperm heads were observed continuously until EGK.V.

The amount of polyspermy seems to be inconsistent. The total number of transmitted sperms varied from 29 to 164,000 per egg among birds, and is positively correlated with egg size (Birkhead et al., 1994). Moreover, the variation of sperm number was shown even in one species, chicken (Lee et al., 2013a). Also, on the insufficient environment in chicken and zebra fish for fertilization such as being inseminated by a sperm-depleted male or having limited opportunity to copulate, females could compensate for the restricted

number of sperm through the oviduct and make them reaching to oocyte more than expected (Hemmings and Birkhead, 2015). However, detailed information about their functions in gene expression during fertilization is unknown up to date.

## 2.2. The beginning of transcriptional activation during embryogenesis

In animals, the embryogenesis is dependent on the well-organized genetic programs upon fertilization. The first round of embryonic development is regulated by the maternal mRNAs and proteins contributed by the cytoplasm in oocyte until the onset of zygotic transcription, called zygotic genome activation (ZGA). The massive transcriptional activation through the ZGA governs the establishment of an individual organism. Also, ZGA has been suggested to be the fundamental mechanism for the change of status from totipotency to pluripotency.

The onset of ZGA is basically regulated by the various factors such as the nucleocytoplasmic (N/C) ratio, maternal clock, and epigenetic regulation (Lee et al., 2014). Also, ZGA occurs successively including first and second waves, which is called minor ZGA and major ZGA in mammals. The first-wave ZGA seems to occur during the fertilization, while the second-wave ZGA shows the different timing of the massive transcriptional activation in various species (Lee et al., 2014).

## 2.3. The methods for identification of *de novo* zygotic transcription

To measure the appearance of ZGA, several experiments such as inhibiting transcription by chemicals or genetic manipulation were performed in animal models. For instance, the treatment of an inhibitor of RNA

polymerase II, alpha-amanitin, halted the cleavage progression in mouse embryos at 2-cell stage (Braude et al., 1979). In *Drosophila*, similar experiment inhibited the cellularization when major ZGA groove occurred (Edgar and Datar, 1996). The researches by plasmid-born reporter gene expression have contributed to the insight of clear ZGA in 1-cell embryos in mouse (Wiekowski et al., 1991, Majumder et al., 1993). Moreover, incorporation of bromo-uridine triphosphate (BrdU) detected the newly RNA synthesis during minor ZGA (Bouniol et al., 1995, Aoki et al., 1997). Also, detecting the phosphorylation of C-terminal domain in RNA polymerase II have been used as a marker of ZGA (Nagai et al., 2015, Lee et al., 2016).

In addition to the above experiment, next-generation sequencing (NGS) technology have been used to distinguish *de novo* ZGA from the maternal contribution, because of the remaining mRNA molecules up to 70% at the peak of zygotic expression in zebrafish (Lee et al., 2013b). After injection of nucleoside analogs such as 4-thio-uridine triphosphate into 1-cell zebrafish embryos, the incorporated and newly expressed transcripts could be selected with biotinylated at the thiol moieties (Heyn et al., 2014). Subsequently, pull down with Illumina sequencing showed the numerous and enriched gene expression by ZGA.

Furthermore, sensitive RNA sequencing (RNA-seq) techniques could detect the definite expression from ZGA. In most cases in *de novo* transcription, both the maternal and paternal alleles are expected to arise in one organism. Thus, the detection of paternal alleles based on the informative single-nucleotide polymorphisms (SNPs) to distinguish allelic expression. Using two different strains in zebrafish and RNA-seq on their hybrid embryos, Harvey et al. found the expressed genes bearing both parental

genotypes by ZGA (Harvey et al., 2013). Also, genome activation were featured in bovine embryos between *Bos t. taurus* cows and *Bos t. indicus* bull (Graf et al., 2014). Finally, the capacity to sequence unspliced pre-mRNA molecules would also detect the ZGA. Using intron signals of primary transcripts from whole-transcriptome RNA-seq, as an alternative to traditional poly (A) + selective library, zebrafish embryos could successfully detect de novo expression from the active zygotic genome in the late blastula stage (Lee et al., 2013b). Thus, depending on NGS technologies, the various evidences of ZGA in chicken could be elucidated at the genome-wide resolution.

### 2.4. First-wave transcriptional activation

The first-wave transcriptional activation, also known as minor ZGA in mammals, were found throughout the species. In mouse, minor ZGA occurs in both parental pronucleus (PN) during fertilization, specifically in the late pronuclear stage (PN4-5), using BrdU (Bouniol et al., 1995). Interestingly, it has been known that the timing and global gene expression levels between paternal and maternal PN are different (Bouniol et al., 1995, Aoki et al., 1997). The transcriptional activation from male PN showed earlier and greater than that of maternal PN. Similar phenomenon was also observed in Sea urchin embryos using radio-labelled uridine (Longo and Kunkle, 1977, Poccia et al., 1985).

These asymmetric contribution to minor ZGA is thought to be caused by the difference between each chromatin properties. During the pronuclear stage, the dynamic changes of its chromatin state and histone modifications drive the transcriptional activation from both paternal and maternal PN (Wu et al., 2016). A unique chromatin state, such as globally permissive shape in actively expressed genes including MERVL and ZFP382 by minor ZGA, were

found. Also, the allelic reprogramming is occurred in histone modification in histone H3 lysine 4 (H3K4) tri-methylation and shown to be established mainly by H3K4 methyltransferase MII3/4 (Aoshima et al., 2015, Zhang et al., 2016).

The first murine zygotic expression was deeply investigated using total RNA-seq (Abe et al., 2015). The highly promiscuous transcripts such as intergenic region-derived and untranslatable mRNAs were abundant at comparably low levels in 1-cell stage in mouse. The authors implied that these irregular gene expression regulation in 1-cell embryos might protect against the precocious gene product from the loosen chromatin states.

In zebrafish, the earliest activated genes were examined by RNA-seq from the only zygotic transcripts by pull-down method (Heyn et al., 2014). At fertilization, 1-cell zebrafish embryo highly activated from the mitochondrial genome. Characteristics of those genes are the relatively short length and evolutionarily younger. They also compared fish, fly, and mouse transcriptome and revealed the different subset of genes from first-wave ZGA.

Even, in plant zygotes, unequal parental genome activation were observed (Anderson et al., 2017). Using hybrid rice zygotes, large-scale transcriptomic dynamics were found soon after fertilization. Also, the highly asymmetric contribution to ZGA were exhibited; the maternal genome activated the genes for basic cellular functions, whereas the paternal genome was restricted to the expression of putative pluripotent factors. Also in maize, ZGA in occurs in zygotes and shows 10% of genome highly activated in dynamic patterns (Chen et al., 2017). Soon after fertilization, transcriptional regulator genes involved in various families are activated. Furthermore, chromatin assembly is strongly modified and suggested to be associated with

ZGA in zygotes. Thus, the profound studies about first-wave transcriptional activation should be conducted for the species-specific features in zygotes, even in avian species.

## 3. Cleavage stages

### 3.1. Asymmetric cellularization and layer increase

Based on the criteria of Eyal-Giladi and Kochav (EGK) (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980), EGK.I is defined as the stage of first cleavage. After that, the rapid and meroblastic cleavage continues for the next 10 hours until EGK.VI. Smaller cells are produced as development progresses, and cell layers are gradually increased until 5 to 6 thickness. During this process, the subgerminal cavity is firstly appeared at EGK.III.

In total, nuclear division cycle is about $12^{th}$ to $14^{th}$ and 8,000 cells were observed on average at EGK.VI. Along with cell division, the mechanism of cell layer increasing were hypothesized that oriented cell division lead to increasing of cellular depth. However, the previous report concluded that the increase in cell layer number was not directly caused by orientation of cell division (Nagai et al., 2015). However, the orientation of division in deep cells showed lower proportion of the angle less than 10° than surface cells (Nagai et al., 2015), suggesting that the unknown mechanism could be involved in this division to some extent.

The avian specific cell division was also examined during cleavage stages. Compared to mammals, the very irregular cleavage pattern were exhibited; asynchronous and asymmetric division. The first two division were shown to be synchronized. But, afterward, the initial cruciform cleavage

generated the four nonpolar embryonic cells (Lee et al., 2013a). Subsequently, cell cleavage from EGK.II to V were progressed with the radial manner. 8-cell staging embryo contains one closed cell and seven open cells, and the central cells were more rapidly divided than peripheral cells (Lee et al., 2013a). However, these phenomenon are justified morphologically up to date.

### 3.2. Second-wave transcriptional activation

The timing of second-wave transcriptional activation, also known as major ZGA, are varied among the vertebrate species (Lee et al., 2014). Until major ZGA, transcription is silenced during the cleavage stages of human, from the 1- to 4-cell stages (Braude et al., 1988, Xue et al., 2013). The major ZGA in mammals occurs in 2-cell and 8-cell stage in human and mouse, respectively (Xue et al., 2013, Yan et al., 2013). In zebrafish and frog, 128-cell stage has been reported as ZGA starts at the maternal-to-zygotic transition (Lee et al., 2014). In chicken, ZGA is solely described as occurring in the late EGK.II to early EGK.III stage based on the phosphorylation of the RNA polymerase II C-terminal domain (Nagai et al., 2015).

The major ZGA is largely depend on maternally inherited transcription factors. The well-known pluripotency-related transcription factors such as NANOG homeobox (NANOG), sex determining region Y-box 2 (SOX2), and POU class 5 homeobox 1 (POU5F1) are largely required for the initiation of ZGA in vertebrates (Lee et al., 2013b, Leichsenring et al., 2013). Ribosome profiling revealed that nanog, sox19b and pou5f1 are the most highly translated transcription factors in ZGA (Lee et al., 2013b). After loss-of-function of these factors, the zygotic genes were failed to be activated and developmental arrest before gastrulation was observed. Among them, Pou5f1 occupies SOX-POU binding sites before ZGA and activates the

zygotic genes (Leichsenring et al., 2013). However, the genetic mechanism, even expression of well-known transcription factors for second-wave transcriptional activation is largely unknown in avian species.

### 3.3. Signaling pathways in asymmetric cell division in early embryos

#### 3.3.1. Notch signaling

Notch signaling is an ancient and evolutionarily conserved intercellular signaling mechanism, which is essential for developmental processes in embryonic and adult tissues including cell differentiation process, anterior-posterior polarity, left-right asymmetry and somitogenesis through the cell-cell interaction regulation (Shi and Stanley, 2006). The Notch genes (Notch1-4) are the receptors stimulated by Delta and Jagged ligands. Their interaction leads to a subsequent proteolytic cleavages releasing the Notch intracellular domain (NICD) from the plasma membrane. NICD translocates to the nucleus and forms a complex with the DNA binding protein CSL (RBPJ), replacing CSL-bound co-repressor complex with histone deacetylase (HDAC). The transcriptional activation complex members, including MAML1 and histone acetyltransferases (HATs), bind to the NICD-CSL complex and lead to induce Notch target genes.

Notch signaling is extensively used for the earliest cell fate decisions. In *Caenorhabditis elegans*, these interactions play the key role in early cell fate determination (Priess, 2005). At the 4-cell stage, the Notch ligand-expressing P2 cell activates Notch signaling in the ABp cell and represses the expression of TBX-37 and TBX-38. Consequently, ABp descendants have the different cellular fate from that of ABa descendants. At the 12-cell stage, the MS cell gives a Notch activation to two of the ABa

descendants in which Notch signaling participates with Tbox genes, which lead to mesodermal fate. Notch signaling will be used for some of their descendants to be cellular fates diversification by the next cell divisions. However, in other species such as sea urchin and zebrafish, WNT signaling in cell fate determination is involved during later stages in blastula and gastrulation, mainly into mesodermal induction.

### 3.3.2. WNT signaling

The WNT signaling is required for basic developmental processes such as controlling cell proliferation, cell fate determination, and body-axis formation during early embryonic development in vertebrates (Hikasa and Sokol, 2013). WNT signaling consists of three different pathways including the beta-catenin-dependent canonical pathway, the planar cell polarity (PCP) pathway, and the WNT/Ca2+ pathway (Niehrs, 2012). In the canonical pathway, the major effector of WNT signaling is the stabilization of cytoplasmic beta-catenin through inhibition of the beta-catenin degradation complex. Then, beta-catenin is entering the nucleus and activates target genes through the interaction with TCF and LEF family transcription factors and DNA binding. The PCP signaling activates the small GTPases, RhoA and RAC1, which ultimately induce the stress kinase JNK and ROCK, and leads to cytoskeleton remodeling and changes in cell adhesion and motility. WNT/Ca2+ signaling is mediated by G proteins and phospholipases and promotes the transient increasing of cytoplasmic calcium subsequently activating the kinases, PKC and CAMKII, and the phosphatase, calcineurin.

Wnt signaling to regulates cell polarity and cleavage patterns during early cell division in various species. During early development, maternally inherited WNT transcript, *Wnt11*, is essential for the early dorso-ventral axis

formation through the canonical pathway in *Xenopus* (Tao et al., 2005). Furthermore, maternal *Wnt5a* and *Wnt11* are regulated by maternal Wnt antagonist, *Dkk1*, in terms of both canonical and non-canonical pathways for embryonic patterning in *Xenopus* (Cha et al., 2008). Recently, the beta-catenin independent non-canonical WNT/STOP signaling pathway regulated by maternal WNT and co-receptors, Lrp6 and Cyclin Y, plays a role in cell cycle progression during early *Xenopus* embryogenesis (Huang et al., 2015). Additionally, small GTPase, RhoA regulates spindle formation and cytokinesis during early embryonic development in *C. elegans* and *Sus scrofa* (Tse et al., 2012, Zhang et al., 2014). Furthermore, the WNT/PCP pathway, through RhoA interaction with the actin cytoskeleton, regulates the orientation of cell division (Castanon et al., 2013). Thus, the downstream pathways of WNT signaling could be involved in the various aspects of early cell division in many species.

### 3.3.3. MAPK signaling

The mitogen-activated protein kinase (MAPK) is a highly conserved cascade that is involved in various cellular functions, mainly in cell proliferation, differentiation, and migration. Among the distinctly regulated groups of MAPKs, extracellular signal-related kinases (ERK)1/2, Jun amino-terminal kinases (JNKs), p38 proteins (p38), activation of ERK1/2 by growth factors depends on the MAPKKK c-Raf and the MAPKK MEK1/2 and leads to cell proliferation.

In sea urchin embryos after fertilization, the MAPK pathway, via cdk2 activation, regulates DNA replication and entry into mitosis (Zhang et al., 2005, Kisielewska et al., 2009). After fertilization, MAPK becomes dephosphorylated form, but small number of ERK-like proteins reactivated

during mitosis. The small fraction of activated MAPK signaling proteins is localized in nuclear region and mitotic poles and spindles. Furthermore, this ERK is dependent on calcium signal mediated by fertilization. Also activation of MAPK promotes cyclinE and cdk2 accumulation in pronuclus to entry into cell division.

## 4. Area pellucia formation period

### 4.1. The first axis formation and lineage segregation in morphological aspects

The second half of intrauterine development includes EGK.VII to EGK.X (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980). The first anterior-posterior axis formation with layer reduction and early lineage segregation occur in the period of area pellucida formation, indicating not basic cell division but the dynamic morphogenesis during the last 10 hours before oviposition. The cell layers in chicken embryos increase rapidly from EGK.I to EGK.VI via cell division, and subsequently decrease to constitute single- and bi-layered EGK.X embryos. Simultaneously, the first lineage specification of the epiblast and hypoblast, and area opaca formation, occur during intrauterine development in the chicken (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980).

The anterior-posterior axis formation is shown to be established by egg tilting and reorienting the yolk by gravity forces in oviduct during this period (Kochav and Eyal-Giladi, 1971). Together with axis determination, layer reduction begins from the posterior side and spreads to all areas except for future area opaca regions.

There are two proposed mechanisms for decrease of cell layers. The first one is cell shedding. This shedding phenomenon is thought to be related to axis formation due to its occurrence in a future posterior to anterior direction. An electron microscopy study revealed that interconnecting fibers, or filophodia, of the lower layer cells in the early chicken embryo were disrupted or retracted during the cell shedding process (Fabian and Eyal-Giladi, 1981). In quail, shedding cells consisting cytoplasmic organelles such as mitochondria, membranous materials, and vesicles and granules were found in the central region of subgerminal yolk using transmission electron microscopy (Andries et al., 1983). The putative shedding cells seem not to be involved in epithelialization and contribution to future blastodermal cells. Cell shedding and formation of area pellucida were shown to be energy consuming, and utilize glycogen reserves accumulated in cells (Kochav et al., 1980). Although an early microscopic study (Kochav et al., 1980) showed cell shedding along the formation of area pellucida, biological significance and mechanism of cell shedding require further detailed elucidation including live imaging. The second one is radial intercalation as an alternative but not exclusive mechanism against cell shedding. Considering rapid layer reduction during these stages, the cellular rearrangement and movement could be also play an important role in layer reduction and epithelization of blastoderms, like as frog or fish embryos (Keller, 2002).

Finally, at EGK.X, the region of area pellucida consists of single- or bi-layered with epiblast and hypoblast and appears more transparent than area opaca which is still remains multilayered. There are two possible mechanisms for the first lineage segregation including predetermination as a salt-and-pepper manner and sorting, and lineage determination after cell layer reduction (Sheng, 2014). The former seems more convincing because of the

previous research in zebra finch (Mak et al., 2015). The differentiation-related transcription factors including *SOX3* as an epiblast marker, and *SOX17*, *GATA4* and *GATA6* as endoderm markers, were examined in the zebra finch embryos staging EGK.VII to VIII (Mak et al., 2015), suggesting molecular determination takes place prior to morphological differentiation in birds. However, the possibility for the specification of cell fate by maternal determinants at a much earlier stages could be demonstrated in avian species (Callebaut, 2005).

## 4.2. Maternal-to-zygotic transition (MZT) is the conversion from maternal contributions to embryonic products

The initial and critical events of early embryonic development are governed by fundamental cellular processes such as transcription and translation. During early development, MZT is the crucial process for generating the individual organism out of the maternal factors. During MZT, the massive induction of zygotic transcription by second-wave (major) ZGA and degradation of maternal mRNAs by zygotic microRNAs (miRNAs) happen in this period. Thus, MZT is composed of transcription including miRNAs from the embryo's own genome and the clearance by its products.

The timing of MZT is dependent on second-wave ZGA, which means the various points of transition between zygotic and maternal among the vertebrates (Tadros and Lipshitz, 2009). The second-wave ZGA is at another time point such as 2-cell in mouse, 8-cell in human, and 128-cell in zebrafish and frog, but the purpose of MZT is the same. In addition to transcriptional aspects, the fluctuation of ribosomal RNA (rRNA) is also the evidence of MZT. In bovine, pre-hatching early embryos showed a lower content of 28S rRNA, indicating the limited potential of translation before

MZT (Gilbert et al., 2009). Also, the transition of maternal and zygotic ribosome pools were observed during MZT in early mouse embryos (Ihara et al., 2011). Thus, transcriptional and translational dynamics is the focal point of MZT.

The highly and evolutionarily conserved miRNAs by the ZGA are key factors for the degradation of maternal RNAs in various species (Svoboda and Flemr, 2010). MiRNAs, which are 18–23 nucleotides in length, bind to the complementary sequences of a target mRNA at the 3′ untranslated region (UTR), and silence it through translational repression or mRNA degradation. After ZGA, the zygotic miRNAs, such as dre-miR-430s in zebrafish, xla-miR-427 in frog and mmu-miR-290s in mouse, were expressed and target the maternally inherited RNAs in different timescale (Svoboda and Flemr, 2010, Yartseva and Giraldez, 2015). Importantly, these miRNAs shared the same AAGUGC motif in the proximal site of matured form. In chicken, gga-miR-302s are predicted to be the ortholog of AAGUGC miRNAs (Svoboda and Flemr, 2010). Also, the miR-302 family is specifically expressed in chicken blastoderm, as miR-290 family in the mouse embryo, suggesting that zygotic miRNA related to MZT also seem to be related to pluripotency.

The mid-blastula transition (MBT) is the term originated from the amphibian development (Tadros and Lipshitz, 2009). Specifically, the approximate time point of mid-blastula stage, after $12^{th}$ cell cycle in frog, was defined as MBT. Here, the cell cycle are getting slower and asynchronous, which is related to embryonic morphogenesis and the activation of zygotic transcription. In both frogs and zebrafish, ZGA is crucial for the progress of gastrulation after the MBT (Newport and Kirschner, 1982, Zamir et al., 1997). According to these characteristics, the MBT is the morphological embodiment

for the MZT in frog and zebrafish.

In chicken, a rapid celluarization takes place during cleavage stages and subsequently, MBT-like phenotype such as prolonged cell cycle around EGK.VI was observed based on cell counting (Park et al., 2006, Nagai et al., 2015). However, the molecular basis in the induction of zygotic developmental program and clearance of maternal factors by post-transcriptional regulators during MZT has not been studied in avian species to date.

### 4.3. Signaling pathways in axis formation and lineage segregation

#### 4.3.1. WNT signaling

In addition to early cell division stages, WNT signaling is involved in later stages. In the preimplantation embryos of mammals, the first cell fate decision occurs in morula stage into trophoectoderm and inner cell mass. WNT antagonists, DKK1, are shown to be leading to differentiation of the trophoectoderm and hypoblast lineages, but inhibiting epiblast formation (Denicol et al., 2014). Also, the aberrant WNT signaling reduces number of trophoectoderm and inner cell mass cells in bovine morula embryos and various WNT signaling-related genes are involved in this stage (Denicol et al., 2013), suggesting WNT signaling system and the downstream activation is essential for early embryonic development and cellular fates.

Differ from cleavage stages, some evidences of WNT signaling involvement in chicken development has been provided. The main effector of canonical WNT signaling, beta-catenin, is firstly localized in late intrauterine stages (Roeser et al., 1999). In EGK.V, beta-catenin is observed in peri-

nuclear region and co-localized with the mitotic spindle apparatus, indicating possible association with the microtubule network. From EGK.VIII, beta-catenin shows nuclear localization in the periphery of the blastoderm, the future area opaca and marginal zone. This suggests that the onset of posterior axis and hypoblast formation could be prepared by beta-catanin dependent canonical WNT signaling prior to oviposition. Furthermore, within WNT ligand family, only *WNT8C* is expressed in posterior marginal zone in post-ovipositional stages (Skromne and Stern, 2001). Its expression is decreasing gradiently posterior to anterior. In addition, WNT antagonists, *CRESCENT* and *DKK1*, seem to block the downstream signaling mediated by *WNT8C*.

### 4.3.2. TGF-beta signaling

The transforming growth factor-beta (TGF-beta) family genes, including TGF-beta, Activin and bone morphogenetic proteins (BMPs), are the cytokines found in various species from worms to mammals. A wide cellular functions such as proliferation, apoptosis, differentiation and migration are mediated by TGF-beta signaling. When TGF-beta family binds to the Type II receptor and recruits Type I, Type II receptor phosphorylates and activates Type I, which in turn phosphorylates receptor-activated Smads such as Smad2 and Smad3. Upon phosphorylation in Smads, they are associated with the co-Smad, Smad4, and form the heteromeric complex translocating into the nucleus. In the nucleus, the complexes bind to target genes with other DNA-binding and coregulatory factors.

TGF-beta/Activin/nodal signaling is required for maintaining pluripotency in human embryonic stem cells (hESCs) (James et al., 2005). Through the SMAD2/3 transducer, TGF-beta signaling control the undifferentiated state and marker expression in hESCs and inner cell mass in

mouse blastocyst. Furthermore in mouse, primed pluripotency mimicking post-implatation embryos are dependent on activing signaling compared to primed state depending on LIF and BMP signaling (Ohgushi and Sasai, 2011). During chicken gastrulation, TGF-beta/Activin signaling is involved in *NANOG* gene expression in epiblast through SMAD2/3 regulation (Shin et al., 2011).

TGF-beta signaling could also cooperate with WNT signaling for the posterior axis and primitive streak formation (Skromne and Stern, 2001). Chicken VG1, also known as GDF1, interacting WNT signaling determines the maginal zone in post-ovipositinal stages. Although the interaction of Wnt and TGF-beta/Activin signaling initiates primitive streak formation after oviposition, their synergism in pre-ovipositional development could be prepared during area pellucida formation period, due to the morphological differentiation to form posterior part in these stages.

## 5. Embryonic dormancy throughout the animal kingdom

### 5.1. Embryonic dormancy across various species

Embryonic dormancy is the suspended development due to the unfavorable environment for normal growth such as temperature, humidity, osmotic pressure, and nutrition. Embryonic dormancy occurs across the animal kingdom (Fenelon et al., 2014), but the detailed regulation and the stage are diverse in each species. The dormancy in early embryos in animal consists of two modes; cold torpor and diapause.

When the development has already begun, cold torpor can slow down the embryo growing periodically (Ruf and Geiser, 2014). Cold torpor

occurs in many species by reducing their energy spending to tolerate short period of stressful events such as cold exposure (Geiser, 2004). On the other hand, embryonic diapause refers that the complex developmental program lead to the conformational changes to slow cell division, reduce metabolism rate and enhance stress tolerance against unfavorable environments (Renfree and Shaw, 2000).

The molecular characteristics about embryonic dormancy have been largely studied in aquatic organisms. The embryos of brine shrimp can withstand the extreme anoxia or hypothermia. When the unfavorable conditions such as freezing or anoxic are faced, the embryos form a cyst, which is produced oviparously with rigid, chitinous shell (Qiu et al., 2007). The encysted embryo show unique gene expression, and mainly the LEA gene contribute the tolerance of the embryo against external stress factors (Toxopeus et al., 2014). In addition, after oviposition of the annual killfish, the embryo enters into the metabolic dormancy status enabling survival without oxygen (Podrabsky and Culpepper, 2012). In diapause fish embryos, ATP contents were greatly decreased, which results in the cell cycle arrest in G0/G1 phase, under the anoxic conditions.

In mammalian species, the implantation of embryo in uterus is suspended by the stimulation of estrogen and progestin on the dormant blastocyst to control the mitotic arrest in diapause state (Lopes et al., 2004). Recently, several signaling pathways is known to be inducers of paused state in mouse. The major nutrient sensor and activator of growth, mechanistic target of rapamycin (mTOR), is related to embryonic dormancy of mouse blastocyst (Bulut-Karslioglu et al., 2016). The partially repressed mTOR signaling induces the reversible diapause of mouse embryos. In addition, FGF

signaling and autophagy affect the formation of multivesicular bodies for the activation of diapause blastcyst (Shin et al., 2017). These studies could be implicating to the fields of reproductive technology, regenerative medicine, etc.

### 5.2. Cold torpor and blastoderm dormancy in avian species

In avian species, the mothers start to brood of eggs at once, after the last egg have been laid and the clutch begins (Winkler and Walters, 1983). In this situation, the firstly laid egg should suffer from the longest exposure to the below temperature for re-initiating further development. Thus, cold torpor occurs abundantly in avian or reptilian species during each clutch period. Upon the beginning of incubation, the embryos in cold torpor can develop normally.

The cold torpor capacity to resume normal development after oviposition is important to not only wild birds but also domestic poultry, because of their egg production. The egg storage is common in poultry industry, and important for obtaining healthy chicks (Gomez-de-Travecedo et al., 2014). In general, the chicken embryo can endure up to 2 weeks at approximately 15°C to 20°C with minimum effects on the hatchability (Fasenko, 2007). After oviposition, the internal temperature of egg drops in several hours to ambient temperature of the hen house (Patterson et al., 2008). The chick embryo in oviposited egg is confront with low temperature and enter into blastoderm dormancy until the optimal development initiate again.

The molecular and cellular characteristics in blastoderm dormancy in chicken is still handful of information during cold torpor. As the cold storage takes longer, the proportion of apoptotic cells in chicken blastoderm were

gradually increased (Hamidu et al., 2010). Also, the expression of genes related to anti-apoptosis and heat shock protein 70 in chicken blastoderm were examined, indicating the blastodermal cells have their own ability to protect themselves against external stresses (Bloom et al., 1998).

Recent report have addressed the extensive molecular studies including transcriptome, the activation of signaling pathways, apoptosis, and cell cycle status (Ko et al., 2017). In this study, to understand the molecular signatures in blastoderm dormancy in chicken, several differentially regulated stress tolerance pathways in the dormant chicken blastoderm were identified. Accordingly, the up-regulation of genes and proteins related to endoplasmic reticulum (ER) stress and stress-activated protein kinase (SAPK) signaling in low temperature stored blastoderms. In addition, the proportion of early apoptotic cells rose dramatically during cold torpor, whereas the proportion of late apoptotic cells was unchanged. Cell cycle was arrested at $G_2$ phase in a DNA damage–independent manner. Interestingly, all events during cold torpor could be reversible to normal states after starting incubation at 37°C.

### 5.3. Species-specific small heat shock proteins in embryonic dormancy

Heat shock proteins (HSPs) are increased in various cellular stresses and function as molecular chaperone binding and inhibiting irreversible protein aggregation or misfolding under stressful conditions (Richter et al., 2010). Among HSP families, small heat shock protein (sHSP) family ranges in size from 12 to 42 kDa, including highly variable N-terminal and C-terminal region and conserved α-crystallin domains (Basha et al., 2012). Monomers of sHSP can interact and bind themselves via α-crystallin domain to form dimers or high molecular oligomers assisted by N-terminal and C-terminal region. Unlike other HSPs, sHSPs function as holdase in the absence of ATP and can

bind different kinds of protein substrates contributing to cell survival (Mymrikov et al., 2011). ATP-independent holdase function of sHSP is important especially in which ATP concentration is low or limited.

sHSPs has been known as co-chaperone components for anti-apoptotic, anti-oxidant, and pro-autophagic functions in cell survival mechanisms (Mymrikov et al., 2011). HSPB1 appear to prevent releasing cytochrome *c* from mitochondria (Paul et al., 2002). HSPB1 interacts with phosphatidylinositol 3-kinase (PI3K), which is activating AKT and subsequently phosphorylating Bax, and results in preventing liberation of cytochrome *c* (Havasi et al., 2008). Additionally, HSPB1 inhibits the apoptosis-signal regulated kinase 1 (ASK1) and JNK pathway leading to Bax activation for cytochrome *c* leakage (Stetler et al., 2009). Moreover, HSPB1 has direct interaction to cytochrome c to prevent apoptosome formation (Bruey et al., 2000). HSPB1 modulate the main effector caspase-3 inhibition through procaspase-3 interaction (Pandey et al., 2000) or Smac inhibition (Chauhan et al., 2003), which in turn prohibit apoptosis. Also, phosphorylated HSPB1 seems to interact with Daxx and inhibit its binding to Fas receptor and Daxx-dependent apoptosis (Charette et al., 2000).

HSPB1 also promotes cell viability under oxidative stress condition. High molecular unphosphorylated oligomers of HSPB1 increases glutathione (Mehlen et al., 1996) through the activation of enzymes for glutathione synthesis, which are crucial for the maintenance of intracellular redox potential (Preville et al., 1999, Escobedo et al., 2004, Arrigo, 2007). HSPB8 is participated in regulation of autophagy stimulation. HSPB8 interacts with Bag3, the stimulator of macroautophagy and promote autophagy to proteolysis of misfolded proteins (Carra et al., 2008). Additionally, their

complex activates phosphorylation of the eIF2alpha through the unknown protein kinase, which results in essential protein synthesis and stimulating macroautophagy (Carra et al., 2009). Thus, sHSPs are capable of protecting environmental stresses against cells in various ways.

In particular, sHSP genes were increased and specific expression in dauer of Caenorhabditis elegans (C. elegans) and cyst of Artemia franciscana (A. franciscana), so called diapause state described in many reports as stress-tolerant and developmental arrest (Ludewig et al., 2004, MacRae, 2016). C. elegans dauer with daf-16 activity undergo a dramatic induction of several sHSPs including hsp16.1, hsp16.49, hsp-12.6, hsp-12.3, hsp-20 and sip-1 (Hsu et al., 2003, McElwee et al., 2004). In case of A. fransciscana, p26, ArHsp21 and ArHsp22 are expressed specifically in diapause, and peaking in cyst whereas not detected when developmental promotion (Villeneuve et al., 2006, Qiu and MacRae, 2008b, Qiu and Macrae, 2008a, King et al., 2013). And even lacking of p26 in cyst showed termination of diapause spontaneously (King and MacRae, 2012). In the case of avian species, although embryonic diapause in avian has not been reported, developmental arrest, so called cold torpor, of earlier eggs' embryo in a single clutch is common at the beginning of incubation, until all eggs in the clutch have been laid (Welty, 1982). However, the expressions and functions of small HSPs in stress-tolerant chicken blastoderm have yet to be investigated.

# CHAPTER 3

# The Transcriptome of Early Chicken Embryo Reveal Signaling Pathways Governing Rapid Asymmetric Cellularization and Lineage Segregation

## 1. Introduction

Avian species have been used as valuable research model for virology, cancer, immunology, developmental biology, and biotechnology (Stern, 2005, Streit et al., 2000, Skromne and Stern, 2001, Torlopp et al., 2014, Han, 2009). Due to their important position in the comparative genomics of vertebrate species and wide relevance across many research fields, Bird10K project (B10K) was recently initiated. Through the B10K project, a phylogenetic hierarchy of avian species, which is the comparative genomics for flight and functional adaptations, was generated (Jarvis et al., 2014, Zhang et al., 2014a, Zhang et al., 2015). Although historically significant, no genome-wide study of the initial and critical events of avian embryogenesis before oviposition has been reported, mainly because of practical difficulties in accessing the pre-oviposited embryos.

Based on the criteria of Eyal-Giladi and Kochav (EGK) (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980), intrauterine embryonic development before oviposition consists of 10 stages, which include separate periods of cleavage and area pellucida formation (Sheng, 2014). After first cleavage at EGK.I, the representative processes, such as rapid asymmetric cellularization and increasing cell layer, take place until EGK.VI. The anterior-posterior axis formation by which tilting and gravity forces reorienting the yolk and forming posterior side, and lineage segregation as a salt-and-pepper manner are shown to initiate during EGK.VII – X at the period of area pellucida formation (Kochav and Eyal-Giladi, 1971, Mak et al., 2015). However, for more than a morphological perspective, investigating the temporal regulation of genes controlling the pathways of fundamental cellular program and signaling is essential to understand the early embryonic development for each species.

Transcriptional regulation of early developmental processes, including zygotic genome activation (ZGA), functional networks, and cellular programming, has already been characterized in representative model organisms, such as mice, cows, frogs and zebrafish, as well as humans, using next-generation sequencing technology (Aanes et al., 2011, Tan et al., 2013, Xue et al., 2013, Yan et al., 2013, Cantone and Fisher, 2013, Graf et al., 2014, Lee et al., 2014), but has not been determined in avian species. Recently, molecular studies on cellularization process, pluripotent state, and germ cell specification in avian species have been enabled by the use of non-invasive methods for the acquisition of intrauterine chick embryos and zebra finch embryos at stages EGK.VI to EGK.VII (Lee et al., 2013, Mak et al., 2015, Nagai et al., 2015). Using this method, we performed RNA sequencing (RNA-seq) in oocytes, zygotes and intrauterine embryos of chicken to comprehensively understand transcriptional changes involved in early embryogenesis in this study. To our knowledge, it is an initiative approach of transcriptomic analysis in avian early embryonic development. Through this study, we sought to fill the gap for comparative genomics among the model organisms by identifying the developmental mechanisms in an avian manner.

## 2. Materials and methods

### *Experimental animals and animal care*

The care and experimental use of chickens were approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1). Chickens were maintained according to a standard management program at the University Animal Farm, Seoul National University, Korea. The procedures for animal management, reproduction and embryo manipulation adhered to the standard operating protocols of our laboratory.

### *Chicken early embryo preparation*

The egg-laying times of the white leghorn (WL) hens were recorded, and intrauterine eggs from EGK stages I−VIII were harvested (Figure 3-1A) using an abdominal massage technique (Lee et al., 2013). Briefly, the abdomen was pushed gently until the shell gland was exposed; the surface of the shell gland expanded when an egg was present for egg shell formation. After expansion of the shell gland surface, massaging was used to move the egg gently towards the cloaca until the intrauterine egg was released. EGK stage X blastoderms were collected from WL hens after oviposition. To collect oocytes and zygotes, WL hens were sacrificed and the follicles were collected. Oocyte and zygote were collected from one hen simultaneously. Because of the little transcriptomic differences between pre- and post-ovulatory oocyte in the previous study (Elis et al., 2008), and the infeasible simultaneous acquisition of post-ovulatory oocyte and zygote from one hen, we isolated the only pre-ovulatory large F1 oocyte. Only the zygote embryos not showing cleavage and located in the magnum were collected within 1 hour after fertilization according to the recorded egg-laying times. All embryos

were classified according to the morphological criteria (Figure 3-2). Shortly after collection, the embryos were separated from the egg using sterilized paper, and the shell membrane and albumen were detached from the yolk. A piece of square filter paper (Whatman, Maidstone, United Kingdom) with the hole in the center was placed over the germinal disc. After cutting around the paper containing the embryo, it was gently turned over and transferred to saline to further remove the yolk and vitelline membrane to allow embryo collection.

***RNA isolation and RNA-seq library preparation***

Total RNA was isolated from early embryos using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The quality of the extracted total RNA was determined using the Trinean DropSense96 system (Trinean, Gentbrugge, Belgium), Ribogreen (Invitrogen) and the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) (Table 3-1). Total RNA was used for construction of cDNA libraries using the TruSeq Stranded Total RNA Sample Preparation kit (Illumina Inc., San Diego, CA, USA). The resulting libraries were subjected to transcriptome analysis using the Illumina Nextseq 500 platform to produce paired 150 bp reads.

***RNA-seq data pre-processing***

Based on the RNA libraries generated, paired-end sequencing (150 bp read length and ~150−200 bp insert size) was performed using the NextSeq500 platform (Illumina). To generate clean reads, Trimmomatic (ver. 0.32) (Bolger et al., 2014) was used. Per-base sequence qualities were checked using FastQC (ver. 0.11.2) (Andrews, 2010) and filtered fastq files.

***Alignment and quantification of mapped reads***

The galGal4 Ensembl genome was used as a reference, and filtered reads were aligned onto this reference genome using Tophat2 (ver. 2.0.12) (Trapnell et al., 2009). SAMtools (ver. 0.1.18.0) (Li et al., 2009) was used to convert SAM to BAM files. To count the mapped reads, HTSeq-count (Anders et al., 2014) was used with the merged gene annotation file (.GTF), with mRNAs derived from Ensembl.

***Detecting differentially expressed genes between consecutive stages***

As some recent statistical research suggests, it is appropriate to consider RNA-seq gene expression data as a negative binomial distribution, rather than a normal distribution, in statistical analyses (Robinson and Oshlack, 2010). Based on this, a negative binomial-based generalized linear model (GLM) was used to detect DEGs between before and after stages. In total, six statistical tests, oocyte vs. zygote, zygote vs. EGK.I, EGK.I vs. EGK.III, EGK.III vs. EGK.VI, EGK.VI vs. EGK.VIII, and EGK.VIII vs. EGK.X, were performed using the edgeR package implemented in R (Robinson et al., 2010). Based on edgeR package, one-way analysis of deviance model was firstly fitted for estimating dispersion on the all samples, as follows:

$$\log(E(Expression_i)) = \mu + Stage_i \qquad (1)$$

,where the Stage represents multiple group containing 7 developmental stages across the 21 RNA-seq samples. Because we are only interested between the two adjacent stages of development, a total of six two-group comparisons (Oocyte-Zygote, Zygote-EGK.I, EGK.I-EGK.III,

EGK.III-EGK.VI, EGK.VI-EGK.VIII, and EGK.VIII-EGK.X) were performed using contrast matrix on the (Eqn. 1) model. Under the null hypothesis, $H_0$: Before stage = After stage , a likelihood ratio test was performed. Significance was indicated at a FDR adjusted $P < 0.05$ (Benjamini and Hochberg, 1995).

***Functional analysis of significantly detected genes***

Biological processes from the GO and KEGG databases were used for functional annotation of the detected genes. An enrichment test (hypergeometric test) was performed for each term using DAVID (Huang et al., 2007); here, $P < 0.05$ was considered to indicate a significantly enriched term. Significantly detected GO and KEGG terms were visualized as a correlation-based network. To visualize the network, a distance-matrix is required, thus gene association matrix was generated (if a biological term associates a gene; 1 or not; 0). Then, correlation coefficients can be calculated based on the generated gene-membership vectors for considering N:M relationships between biological terms and genes (estimated correlation is always positive value becaue the vector contains only 0 or 1). A correlation value calculated close to 1 between two biological terms means that two terms are co-occurrence and the closer to 0 represents highly independence. From these correlation-based distance matrixes, a network analysis was performed using the qgraph package in R with spring layout for placement based on centrality.

***In situ* hybridization**

To make hybridization probes, the total RNA from each blastodermal stage was reverse transcribed, and the cDNA was amplified using the primers

of chicken *WNT4*, *WNT6*, *NOTCH1*, and *NOTCH2* (Table 3-14). The PCR products of the correct size were cloned with the pGEM-T Easy Vector System (Promega, Madison, WI, USA). After sequence verification, the recombinant plasmids containing the genes of interest were amplified with T7 (5'-TGTAATACGACTCACTATAGGG-3') and SP6 (5'-CTATTTAGGTGACACTATAGAAT-3') specific primers to prepare the templates for labeling with hybridization probes. Digoxigenin (DIG)-labeled sense and antisense hybridization probes of each gene were transcribed *in vitro* using the DIG RNA Labeling Kit (Roche Diagnostics, Basel, Switzerland). Whole mount *in situ* hybridization was performed following the standard published protocol for chickens. In addition, intrauterine embryos were embedded in paraffin and sectioned at 10 μm on a HM 355S automatic microtome (Thermo Fisher Scientific, Inc., Waltham, MA, USA). After deparaffinization, rehydration, and antigen retrieval, each slide was mounted with Vectashield Antifade Mounting Medium with DAPI (Vector Laboratories, Burlingame, CA, USA). The whole mount embryos and sections were observed and imaged under a Ti-U fluorescence microscope (Nikon, Tokyo, Japan).

***Correlation based co-expressed network analysis of coding genes***

In the co-expression analysis, a total of 42 genes were used including representative protein coding genes and transcription factors that are known to be important in early embryonic developmental of other model organisms (Table 3-12). First, mRNAs included in four representative signaling pathways (Notch, MAPK, Wnt and TGF-beta signaling pathways) were used. Second, the significant transcription factors (ETS1, ETS2, c-myb, NANOG, POUV, SOX2 and MYC) were included. Based on gene expression matrix of

these 42 genes, pairwise correlation tests were performed and correlation coefficients were calculated of all combination of those genes. Based on correlation matrix (42 x 42), the network plot was generated using the qgraph package in R with the spring layout for considering clustering patterns of co-expressed genes. In this plot, only significant relationships (FDR adjusted $P < 0.01$ from the pairwise correlation test) were visualized as edges to identify relationship more conservative. Finally, node size was determined based on the centrality of each gene (# of significant connections at FDR adjusted $P < 0.01$).

***Species comparison of early embryonic gene expression patterns in humans, mice, zebrafish, and chickens***

In various analyses, early embryonic transcriptomic changes in chickens were identified. Although early embryonic transcriptomic changes have been investigated in humans, mice, and zebrafish, no such reports have been conducted in chickens. Thus, one of our aims was to map the developmental stages among different species based on the gene expression patterns identified in this study. For the species comparison, early embryo RNA-seq data of other species were obtained from the public transcriptome database (GSE44183 for human and mice; ERP001280 for zebrafish). In this series, human (oocyte, pronucleus, zygote, 2-cell, 4-cell, 8-cell and morula) samples and mouse (oocyte, pronucleus, 2-cell, 4-cell, 8-cell and morula) samples, and zebrafish (2cell, 64cell, 3.3hpf, 6hpf, and 9hpf) samples were obtained. These obtained raw sequenced files (.fastq) were processed in the same manner as Chicken RNA-seq data pipeline to measure gene expression levels (Quality control, alignment, and quantification methods are same, but reference genome and gene-annotation files are different). In case of reference

genome, genome builds of the hg19, mm9, and zv10 were used for human, mouse, and zebrafish, respectively. For the between species comparison, RPKM normalisation was performed to account for the different gene lengths in each species. In addition, 1:1:1 orthologous genes among human, mouse and chicken were identified using Ensembl Biomart tool to match different gene annotations of three species. As a result, three types of gene expression matrices (including 9,810 genes) were generated for human, mouse and chicken, respectively. In the case of zebrafish, which is evolutionarily distant, only 1:1 orthologous genes between zebrafish and chicken were considered independently. A total 6,221 Ensembl genes were identified as 1:1 orthologous genes in two species. Based on these matrixes, we attempted to calculate relative similarities using the following formula:

$$= \left[\frac{max_{new} - min_{new}}{max_{old} - min_{old}} * \left\{\left(1 - \frac{6\sum rg(X_i) - rg(Y_i)^2}{n(n^2-1)}\right) - max_{old}\right\} + max_{new}\right] / 100$$

$$= \left[100 * \left\{\left(1 - \frac{6\sum rg(X_i) - rg(Y_i)^2}{n(n^2-1)}\right) - 1\right\} + 100\right] \frac{1}{100} \qquad (2)$$

,where rg(X) and rg(Y) are rank values. The score of relative similarity represents extended Spearman's correlation coefficients ([-1 to 1] to [0 to 1] scale). "0" and "1" values, respectively, represent the lowest and highest similarities given pairwise similarities. This statistic was used to identify transition points of similarities in another species given the similarity of all transcripts of other species based on the correlation in different developmental stages of different species. By linear transforming (stretching) the correlations of the whole gene expression levels to a range of values from

0 to 1 given the general similarities of the other species mapping the developmental process, this statistic makes it possible to identify transition points (If there are transition points like as in other species, this statistic will show notable gap).

## 3. Results

### *Acquisition of early chicken embryos and processing of RNA-seq data*

According to morphological staging by Eyal-Giladi and Kochav (EGK) (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980), we chose following stages in shell gland; EGK.I for first cleavage, EGK.III for ZGA, EGK.VI for last stage of cleavage period, EGK.VIII for area pellucida formation in progress, and EGK.X at oviposition (Figure 3-1A). In addition, we also chose oocyte in ovary and zygote in magnum for the changes upon fertilization (Figure 3-1A). Intrauterine eggs were retrieved from oviposition time-checked hens by non-invasive collection (Figure 3-1A and see Materials and methods). In total, 137 of early embryos from oocyte to EGK.X stage blastoderms (15 oocytes from ovary, 9 zygotes from magnum, and 17 EGK.I, 19 EGK.III, 20 EGK.VI, 27 EGK.VIII and 30 EGK.X from shell gland in hen) were collected according to the EGK morphological criteria (Figure 3-2). An average of 3−10 embryos were randomly pooled at each stage, and the total RNA amount per embryo was assessed in each pooled sample (Table 3-1). A single embryo contained 1,457 ng total RNA, on average, through intrauterine development. Based on the total RNA, RNA sequencing was performed using a total of 21 pooled samples (three biological replications per stage from oocyte to EGK.X) on the Illumina NextSeq500 system. As a result, 58,931,937 paired-end reads were available after adapter sequence trimming, and the poor-quality reads were filtered out (Table 3-2) from the originally generated 1,279 million reads (386 Gb). Finally, 35,291,134 paired-end reads were mapped to the Ensembl chicken reference genome (galGal4; Table 3-3), and those reads were mapped to 22,010 genes, after filtering out non-mapped genes across all samples.

***Transcriptomic dynamics representing two waves of ZGA in early chicken development***

Based on quantified gene expression, the relationships among the samples and stages were investigated using multidimensional scaling (MDS) methods (Figure 3-1B). From a stage standpoint, most intrauterine developmental stages were clearly distinguishable, but three stages—zygote, EGK.I and EGK.III—were not. Statistical tests were performed to statistically identify differentially expressed genes (DEGs) between adjacent stages according to the developmental process (i.e. Oocyte-Zygote, Zygote-EGK.I, EGK.I-EGK.III, EGK.III-EGK.VI, EGK.VI-EGK.VIII, and EGK.VIII-EGK.X). As a result, there was no dynamic change among these developmental stages (EGK.I and EGK.III) at 5% significance level with false discovery rate (FDR) multiple testing adjustment, whereas large numbers of DEGs were observed in other consecutive stages (Figure 3-1C).

Among the six consecutive stage comparisons, the greatest changes in gene expression levels (4,819 upregulated and 4,129 downregulated DEGs) were observed between the oocyte and zygote stages (Figure 3-1C). In most cases, equivalent numbers of up- and downregulated genes were observed; however, the significant DEGs between EGK.III and EGK.VI were skewed towards upregulation (1,287 upregulated vs. 13 downregulated). These results may indicate that there are two waves of ZGA, from oocyte to zygote and from EGK.III to EGK.VI, designated as the first (1$^{st}$ ZGA) and second waves of ZGA (2$^{nd}$ ZGA), according to previous reports in chicken and minor and major ZGA in mammals (Lee et al., 2013, Xue et al., 2013, Nagai et al., 2015).

***Developmental network and signaling pathways of cell division after***

***fertilization in chicken***

To identify the functional characteristics of the DEGs between consecutive stages (i.e. Oocyte-Zygote, Zygote-EGK.I, EGK.I-EGK.III, EGK.III-EGK.VI, EGK.VI-EGK.VIII, and EGK.VIII-EGK.X), an enrichment test was performed based on the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases in each of significant genes (FDR adjusted $P < 0.05$). Each of the significantly detected genes in each comparison of consecutive stages were classified as up-regulated and down-regulated genes and analyzed for each. From these analyses, we expected to identify transition-specific and pattern-enriched functions using these gene lists, respectively. As a result, diverse transition-specific (Table 3-4~3-7) were identified (enrichment test P-value $< 0.05$). Because only four transition points, oocyte to zygote, EGK.III to EGK.VI, EGK.VI to EGK.VIII, and EGK.VIII to EGK.X, were observed among the seven developmental stages, only their corresponding DEG lists were analyzed further. The large number of DEGs among the four transitions (Figure 3-1B) produced a broad spectrum of functional terms in our enrichment analysis.

Next, functional term clustering analysis was performed using significantly detected GO and KEGG terms derived from DAVID analysis for considering N:M relationships between biological terms and genes. Based on the significant terms and their associated gene information (See detail in Materials and methods), correlation was calculated to estimate co-occurrence and co-relationships between biological terms (FDR adjusted P derived from correlation test $< 0.01$ was considered as significant relationship). Such identified relationships were visualized as networks to demonstrate the co-occurrence terms.

In the results, many biological terms were observed between the oocyte and zygote stages (Figure 3-3A), which is the 1st ZGA after fertilization, exhibiting the most dynamic changes during early development. Of the observed terms, phosphate metabolic process (GO:0006796), histone acetylation (GO:0016573), Cell cycle (gga04110), DNA replication (gga03030), MAPK signaling pathway (gga04010), regulation of small GTPase-mediated signal transduction (GO:0051056), Wnt signaling pathway (gga04310) and Notch signaling pathway (gga04330) were identified as upregulated. Transmembrane transport (GO:0055085), apoptosis (GO:0006915) and ribosome (gga03010) were downregulated. These results indicated that the genetic network induced by the 1st ZGA may sustain and regulate through the cleavage period until EGK.VI.

Particularly, we deeply investigated the up-regulated functional terms and signaling pathways. Many of cyclins and CDKs and the representative S-phase marker, *PCNA*, were expressed and grouped into Cell cycle and DNA replication (Figure 3-3B; Table 3-8). Also, MAPK signaling pathway genes such as RAS, RAF and several MAP kinases were activated (Figure 3-3B; Table 3-8). In addition, the gene sets of GTPase cycle regulators in regulation of small GTPase mediated signal transduction, the ligands and receptors for Wnt signaling pathway, the downstream effector of Wnt/planar cell polarity (PCP) pathway and Notch signaling pathway-related genes were observed (Figure 3-3C; Table 3-9). We examined the spatial expression of Wnt and Notch signaling-related genes including *WNT4*, *WNT6*, *NOTCH1*, and *NOTCH2* during cleavage stages by *in situ* hybridization (Figure 3-4). As a result, we identified *WNT4*, *WNT6*, and *NOTCH1* were expressed from EGK.I to EGK.VI, and expression of *NOTCH2* were in EGK.I and EGK.III (Figure 3-4A, B). Furthermore, asymmetric expression of these genes were observed

in EGK.III (Figure 3-4C). These biological terms seem to be related with the rapid asymmetric cellularization during cleavage period.

***The definite functional signatures for morphological changes from EGK.VI stage***

Contrary to previous stages, the distinct programs were observed from EGK.VI to EGK.X stages. The functional term networks were characterized and shown in Figure 3-5A-C. By the 2$^{nd}$ ZGA, between EGK.III and EGK.VI, the up-regulated signaling pathways, such as Wnt signaling pathway (gga04310) and TGF-beta signaling pathway (gga04350) were clearly observed (Figure 3-5A). The down-regulated functional terms which are genetic programs of cleavage period such as phosphate metabolic process, histone acetylation, MAPK signaling pathway, regulation of small GTPase-mediated signal transduction, and Notch signaling pathway initially appeared from EGK.VIII to EGK.X (Figure 3-5A-C).

Notably, the regulators of Wnt and TGF-beta signaling including various ligand, inhibitor, receptors, and downstream mediators were found (Figure 3-5D; Table 3-10). Specifically, Wnt signaling-related genes are grouped into three clusters based on expression level (*WNT7A*, *WNT7B*, *WNT9B*, and *WNT3* in one group, *FZD7* in second, and *WNT11*, *WNT3A*, *WNT5A*, *WNT9A*, *DKK1*, *WNT8C*, *WNT8B*, etc. in the third), but we could not find the functional categorization among the ligands in each cluster because canonical (*WMT3*, *WNT3A*, *WNT7B*, *WNT8C*, *WNT8B*), and non-canonical (*WNT5A*, *WNT11*, *WNT7A*) ligand were mixed. These signaling pathways were continuously increased and maintained until EGK.X, indicating that the candidate signaling pathways for the beginning of area pellucida formation after EGK.VI (Figure 3-5A-C). Furthermore, several terms of biological

processes which could be related to lineage specification during intrauterine development were observed, such as neuron development (GO:0048666), heart development (GO:0007507), lung development (GO:0030324) and ECM−receptor interactions (gga04512), which seemed to increase continuously from EGK.VI to X (Figure 3-5E; Table 3-11). These regulators which are distinct with those of cleavage stages seem to govern area pellucida formation period.

***Multi-species comparison among vertebrates including pre-oviposited chicken embryos, pre-implantation human, mouse embryos, and early zebrafish embryos***

The co-expression network analysis based on correlation was performed using gene expressions of protein-encoding genes that are known to be important in early embryonic development, including representative signaling genes which are functioning as ligands, receptors, DNA-binding factors, and kinases of the Notch, MAPK and Wnt, and TGF-beta signaling pathways, along with well-known transcription factor genes were analyzed (Figure 3-6A; Table 3-12). In the co-expression analysis, result of the pairwise correlation tests was employed and 1% significance levels after FDR multiple testing adjustments were considered as co-expressed relationships between two genes. As a result, the network showed a distinct structure between the cleavage and area pellucida formation stages. For instance, all Notch signaling pathway genes, MAPK signaling genes except *MAP2K2*, five Wnt ligands and receptors comprising *WNT4*, *WNT6*, *FZD1*, *RYK*, *CCNYL1* and *ETS2* were enriched only in the right subnetwork; this indicates significant functions in rapid cellularization. In contrast, pluripotency markers, such as *NANOG*, *POUV*, *MYC* and *SOX2,* other five Wnt ligands and receptors including

*WNT5B*, *WNT8C*, *DKK1*, *FZD7* and *FZD10* were positively enriched in the left subnetwork. Regarding the functional characteristics of the genes contained in this subnetwork, they play key roles in area pellucida formation and MZT. These genes are thus candidates of the clearly divided genetic network for the cleavage and area pellucida formation period, during early chicken embryogenesis.

Because our transcriptome analysis showed several signaling pathways such as Notch, MAPK, Wnt, and TGF-beta in the definite two periods in chicken early development, we presumed that gene expression in other well-known species would be different because of the difference in ontogeny between aves and mammals. Additionally, we considered that the counterpart developmental stages in other amniotes needed to be compared with those in the chicken embryo. To address these issues, we performed a comparative analysis of gene expression in pre-implantation embryos in humans and mice (GSE44183 in GEO database) (Xue et al., 2013). In those species, gene expression from the oocyte to the morula stages (oocyte, pronucleus, zygote, 2-cell, 4-cell, 8-cell and morula) was assessed as the preimplantation developmental stages. For the multi-species comparison, 1:1:1 orthologous genes were defined using the Ensembl database. Next, a reads per kilobase per million mapped reads (RPKM) normalization method was used to consider different gene lengths among the species. Finally, a relative similarity measure was used to estimate similarity among the various developmental stages (Figure 3-6B and see details in MATERIALS AND METHODS). Relative similarities tended to increase gradually with developmental progression in the three species. In the comparison between human and mouse, the similar pattern in relative similarity with the embryonic development were shown (Figure 3-7). In addition, minor and major ZGAs

are processed in humans (zygote and 8-cell) and mice (zygote and 2-cell) (Lee et al., 2014), which is consistent with our relative similarity analysis results. Furthermore, the three observed transitions in chicken species were closely correlated with our observations in the transcriptomic analysis. In addition to mammals, we compared the transcriptomes of early embryos in zebrafish (2cell, 64cell, 3.3hpf, 6hpf, and 9hpf) with those of chicken. After identification of 1:1 orthologues between chicken and zebrafish, the relative similarity were examined (Figure 3-8). As in the case of mammals, the values with zebrafish embryo also seemed to increase continuously as the development progress and sharply with ZGA.

In addition to the relative similarity analysis among the species, we compared orthologous gene expression related to cell signaling, regarding to Figure 3-6A during representative stages: oocyte, 1-cell zygote or pronucleus, post-major ZGA (EGK.VI in chicken, 8-cell in human, and 4-cell in mouse), and post-MZT (EGK.X in chicken and Morula in human and mouse) (Figure 3-9; Table 3-13). From these analyses, Notch signaling genes such as *DLL4*, *NOTCH1*, *NOTCH2*, and *RBPJ* were strongly expressed in chicken during all stages, but not in human and mouse (Figure 3-9A). Wnt ligands, including *WNT6* for cleavage and *WNT5B* for area pellucida formation, were also significantly expressed in chicken compared to mammals (Figure 3-9B). MAP kinases, *MAPK1*, *MAP2K3*, and *MAPK14*, which are shown in low expression during early cleavage in mammals, expressed predominantly in cleavage period (Figure 3-9C). TGF-beta family, *TGFB3* were specific to area pellucida formation period, but lower in later stages in mammals (Figure 3-9D). That said, these results indicate that some genes related to signaling such as Notch, MAPK, Wnt and TGF-beta are differentially expressed between chicken and mammals during early developmental program.

Indeed, all of the evidence, including the significantly detected genes and functional terms during the zygote and EGK.VI stages, and zygote and EGK.VIII stages, is illustrated in Figure 3-10. Early embryos in chickens, humans and mice exhibit parallel developmental processes such as cell cycle, transcription, and translation, according to the previous study (Xue et al., 2013) and transcriptomic analysis in this study, despite differences in ontogenetic physiology and expression of signaling pathway-related genes between aves and mammals.

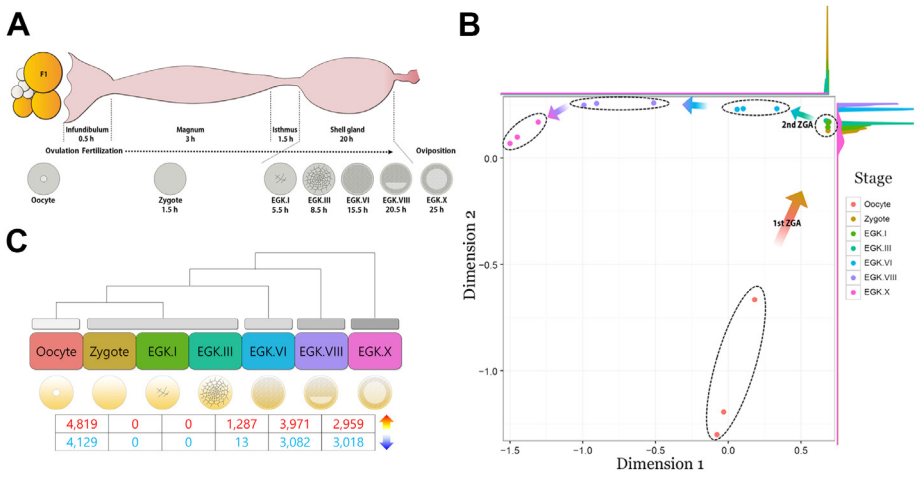


**Figure 3-1. RNA-seq analysis of early chicken embryos.** (A) Schematic of early development and embryo acquisition in chicken. Intrauterine embryos from EGK.I, showing first cleavage, to EGK.VIII, representing area pellucida formation and cell layer reduction in progress from posterior (shaded light gray), were collected from shell gland. EGK.X with fully area pellucida (light gray) and area opaca (outline in dark gray) were obtained after oviposition. Germinal vesicle (round circle in center) oocyte in ovary and zygote without any cleavage in magnum were also retrieved according to egg-laying time. h, staying duration for each part of oviduct and hours after fertilization for each stage of embryos. (B) Multidimensional scaling (MDS) plot, based on gene expression of the whole transcriptome in pre-oviposited chicken embryos. Each point represents an RNA-seq sample, and the dotted circle represents clustered samples. (C) Significantly up- and downregulated genes between consecutive stages (FDR adjusted $P < 0.05$).

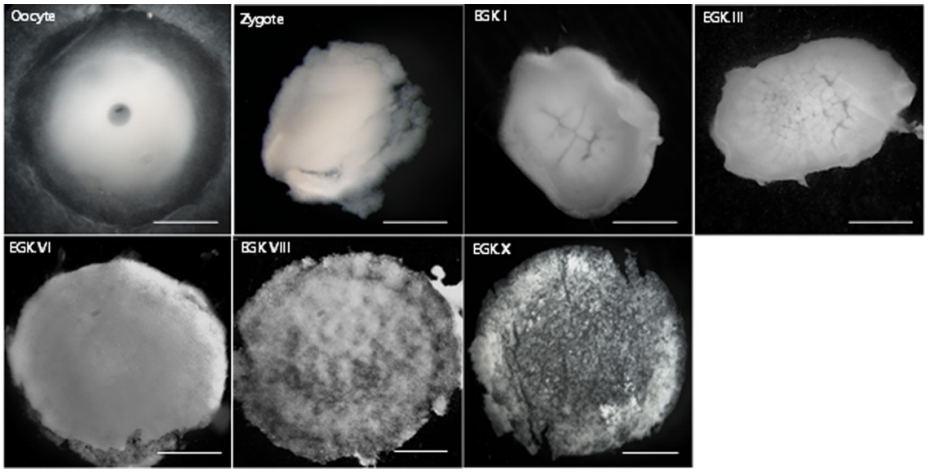


**Figure 3-2. Representative images of early embryos from oocyte to EGK.X in chicken used for RNA-seq.** All embryos following the morphological criteria of Eyal-Giladi and Kochav were used in RNA-seq. Scale bars, 1000μm.

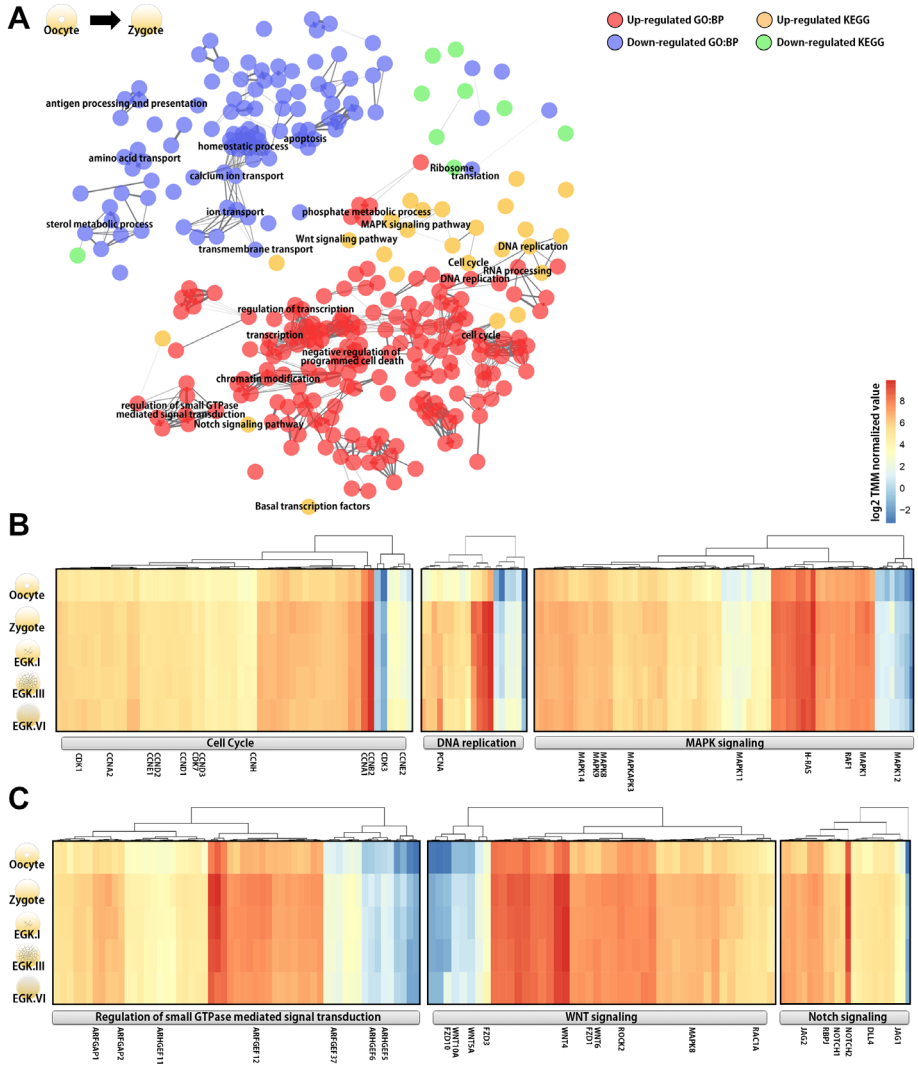


**Figure 3-3. Functional network of significantly expressed genes between oocyte and zygote stage during cleavage period.** (A) Functionally clustered network using significantly detected biological terms, such as biological processes and pathways, from the GO and KEGG databases between oocyte and zygote stage. Each node represents a biological term, and an edge represents a significant correlation between two terms (FDR adjusted $P < 0.01$). (B) The expression patterns of the gene sets

which might be related to rapid cellularization after fertilization. (C) Heatmap showing the expressions of the signaling pathways which could be involved in cell polarity and fate determination during early cleavage period.

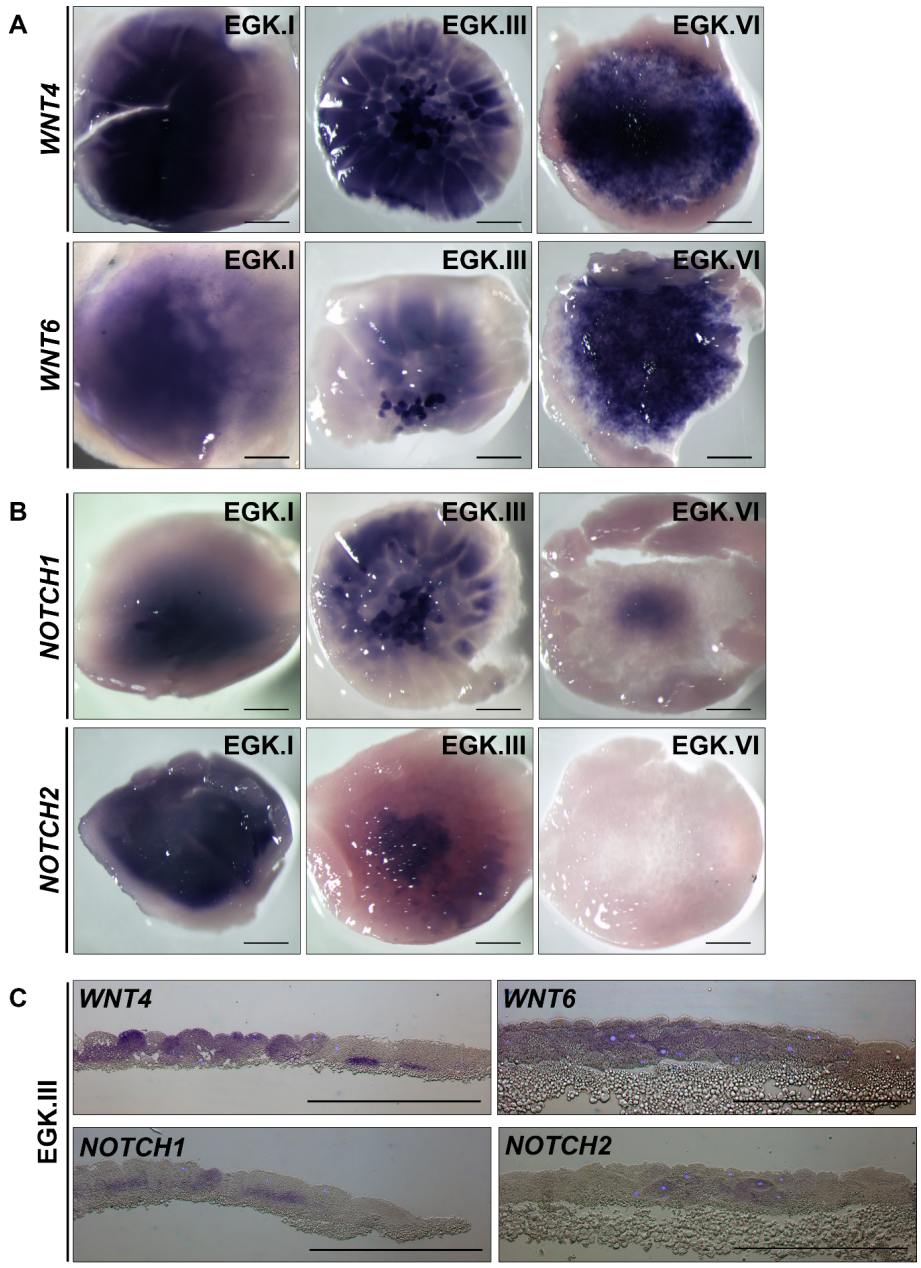


**Figure 3-4. *In situ* hybridization of WNT and Notch signaling-related genes during cleavage stages.** Dorsal views of whole-mount *in situ* hybridization of *WNT4*, *WNT6* (A), *NOTCH1* and *NOTCH2* (B) from EGK.I to EGK.VI. (C) Longitudinal section views of asymmetric expressing *WNT4*, *WNT6*, *NOTCH1*, and *NOTCH2* in

EGK.III. Scale bars = 500 µm.

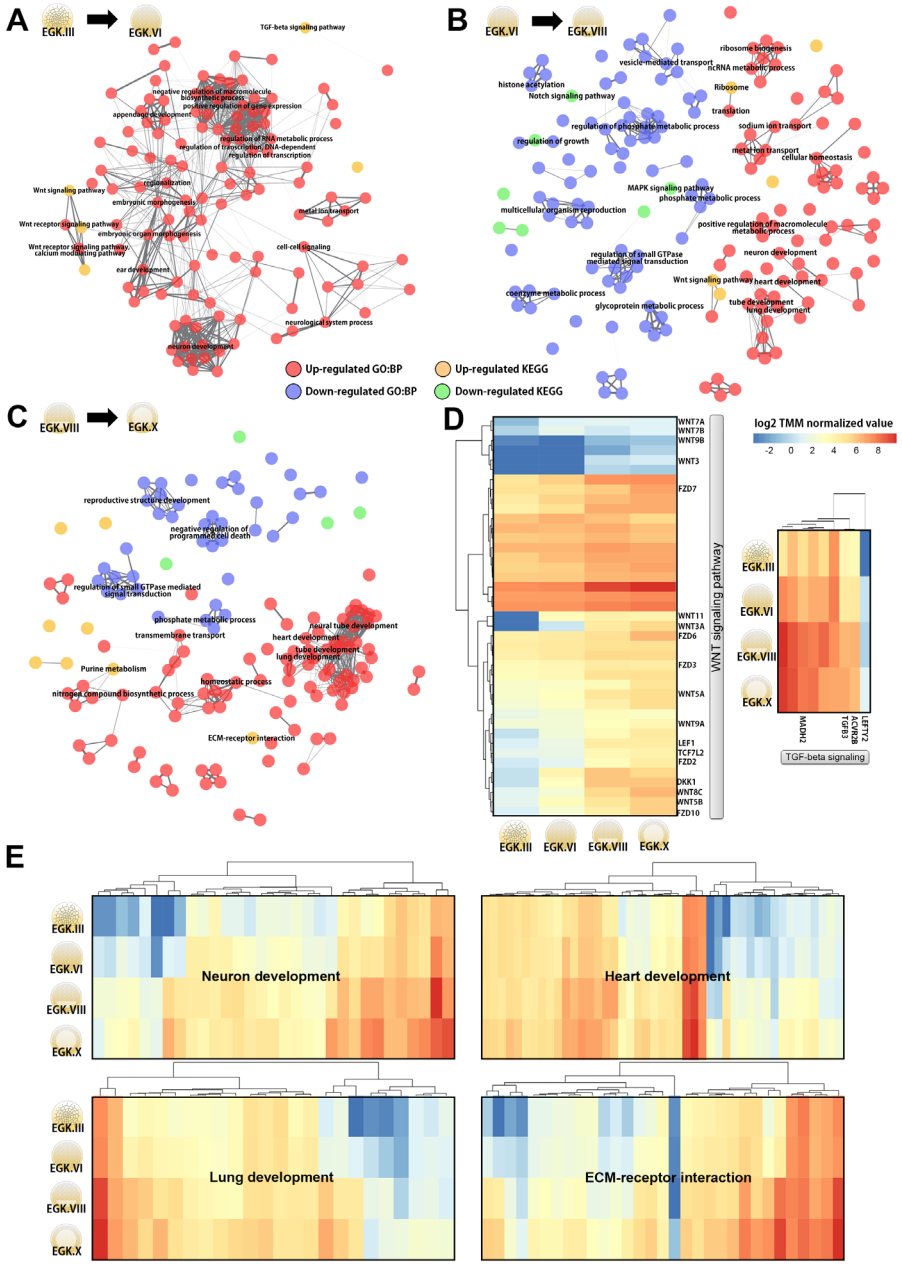


**Figure 3-5. Functional network for the area pellucida formation period after EGK.VI stage.** (A-C) Functionally clustered network using significantly detected biological terms between EGK.III and EGK.VI (A), EGK.VI and EGK.VIII (B), and

EGK.VIII and EGK.X stages (C). Each node represents a biological term, and an edge represents a significant correlation between two terms (FDR adjusted $P < 0.01$). (D) The expression of the gene sets which could be associated with the anterior-posterior axis formation. (E) The expression of the terms which might be involved in lineage segregation and morphogenesis during area pellucida formation period.

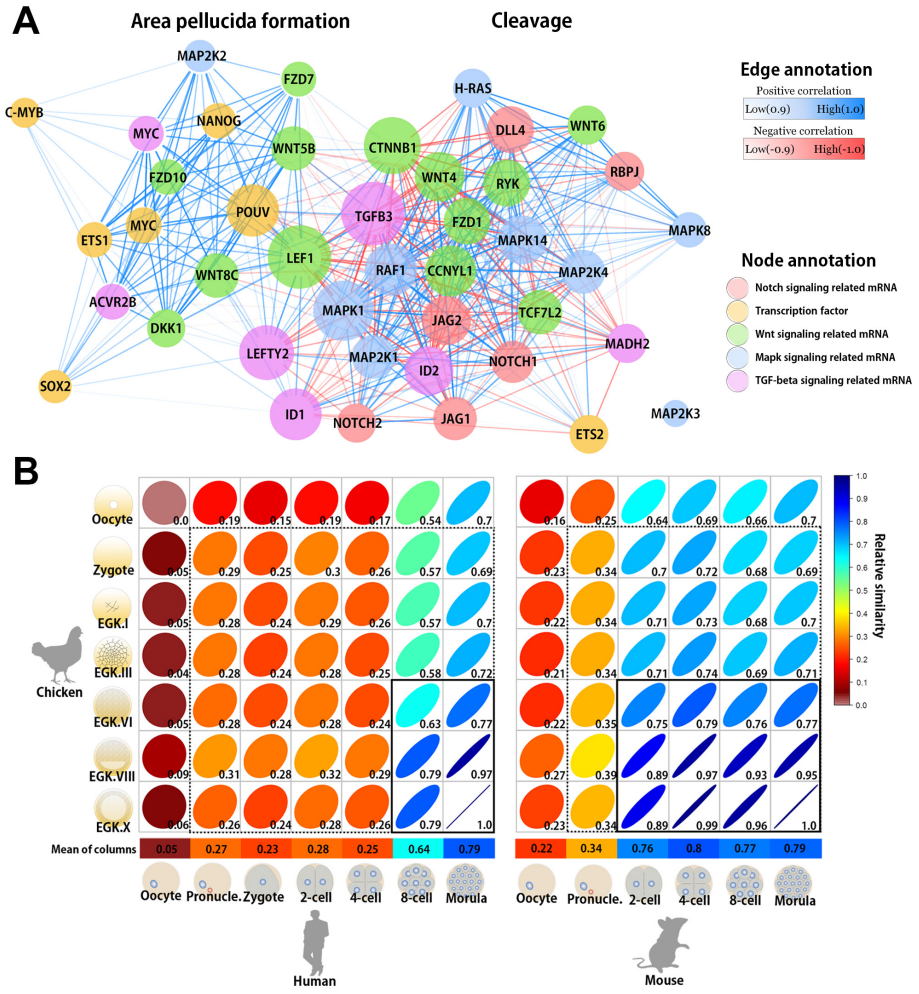


**Figure 3-6. The collective network of distinct developmental program and comparison of transcriptomic profiles between pre-oviposited chicken embryos and early mammalian embryos.** (A) The mRNA co-expression network, composed of diverse transcriptomes, includes representative signaling pathway-related genes and well-known transcription factors. The node size represents centrality, determined based on the significantly detected relationships (FDR adjusted P < 0.01). (B) Heat map obtained from a comparison between early chicken and early human and mouse

embryos. The relative similarity measure was used, based on gene expression in the transcriptome. The color from red to blue shows the higher of relative similarity between stages. The *x*-axis represents the early developmental stages of human or mouse, and the *y*-axis represents the profiling of the early developmental stages of chicken. The dotted and solid boxes, respectively, indicate the first and second transitions in chicken versus the other two species.

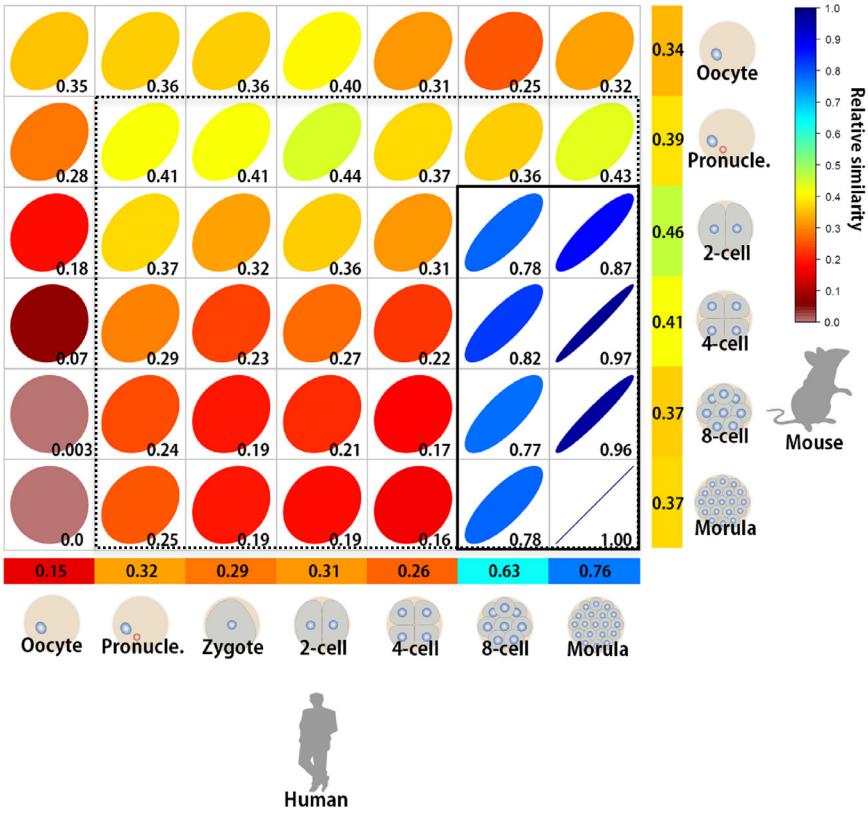


**Figure 3-7. Heatmap obtained from a comparison between early human and mouse embryos.** The relative similarity measure was used, based on the transcriptomes. The *x*-axis represents the early developmental stages of human, and the *y*-axis represents the early developmental stages of mouse. The dotted and solid boxes, respectively, indicate the first and second transitions between two species.

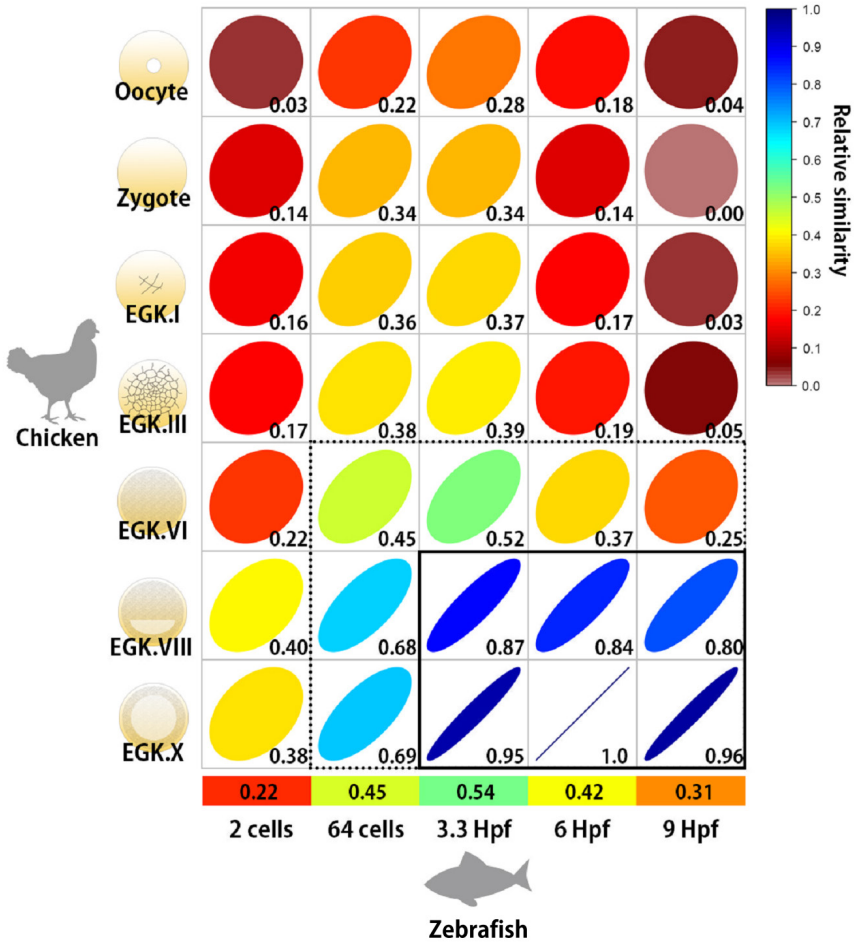


**Figure 3-8. The comparison of transcriptomic profiles between pre-oviposited chicken embryos and early zebrafish embryos.** The relative similarity measure was used, based on gene expression in the transcriptome. The x-axis represents the early developmental stages of zebrafish, and the y-axis represents of the early developmental stages of chicken. The dotted and solid boxes, respectively, indicate the first and second transitions in chicken versus zebrafish.

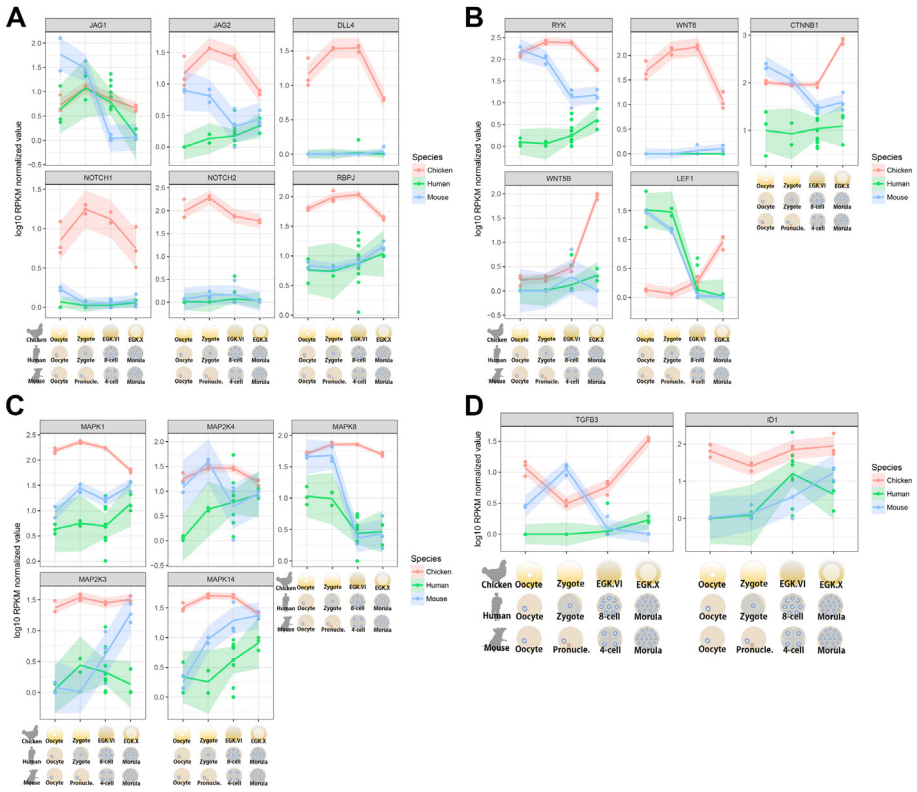


**Figure 3-9. The comparison of orthologous genes during representative stages among chicken, human, and mouse.** (A-D) Gene expression related to cell signaling including Notch (A); Wnt (B); MAPK (C); and TGF-beta (D). Representative stages including oocyte, 1-cell zygote or pronucleus, post-major ZGA (EGK.VI in chicken, 8-cell in human, and 4-cell in mouse), and post-MZT (EGK.X in chicken and Morula in human and mouse) were compared. Normalized RPKM values of genes in each replicate were colored dot and mean values were shaded line among analogous stages in chicken (red), human (green), and mouse (blue).

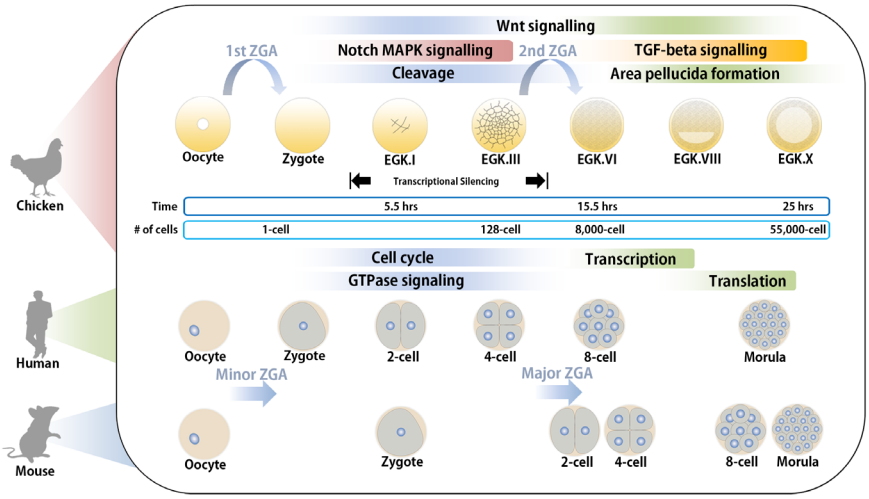


**Figure 3-10. Summary and complementation of early chicken embryo developmental processes with those of human and mouse.** The parallel stages among three species including chicken, human and mouse were mapped based on the analysis of functional terms and developmental processes including ZGA. Signaling pathways such as Notch, MAPK, Wnt and TGF-beta signaling were shown to be avian-specific. The up-regulated genes by 1$^{st}$ ZGA and 2$^{nd}$ ZGA may be involved in regulating Cleavage and Area pellucida formation period, respectively.

**Table 3-1. Quality statistics of RNA sample for RNA-seq**

| No. | Sample ID. | Ribogreen | Volume | Amount | Bioanalyzer | | No. of Embryos | Total RNA per embryo |
|---|---|---|---|---|---|---|---|---|
| | | Con. (ng/ul) | (ul) | (ng) | Ratio (28S:18S) | RIN | | (ng) |
| 1 | Oocyte_1 | 105.5 | 50 | 15827.6 | 1.4 | 8.5 | 5 | 1055.2 |
| 2 | Oocyte_2 | 185.5 | 50 | 27819.9 | 1.5 | 9.3 | 5 | 1854.7 |
| 3 | Oocyte_3 | 167.1 | 50 | 25059.1 | 2 | 9.2 | 5 | 1670.6 |
| 4 | Zygote_2 | 80.3 | 50 | 4016.9 | 0.8 | 7.9 | 3 | 1339 |
| 5 | Zygote_3 | 108.5 | 50 | 5425.8 | 0.1 | 6.5 | 3 | 1808.6 |
| 6 | Zygote_5 | 174.7 | 25 | 4367.4 | 0.2 | 6.6 | 3 | 1455.8 |
| 7 | EGK.I_2 | 174 | 50 | 8697.5 | 0.1 | 6.7 | 6 | 1449.6 |
| 8 | EGK.I_4 | 201.3 | 50 | 10066.6 | 0.8 | 7.2 | 5 | 2013.3 |
| 9 | EGK.I_6 | 141.3 | 50 | 7065.6 | 0.6 | 7.5 | 6 | 1177.6 |
| 10 | EGK.III_2 | 197.1 | 50 | 9855.9 | 0.1 | 7.6 | 6 | 1642.7 |
| 11 | EGK.III_5 | 178.3 | 50 | 8916.9 | 0.8 | 6.5 | 6 | 1486.2 |
| 12 | EGK.III_6 | 188.2 | 50 | 9411.6 | 0.9 | 7.4 | 6 | 1568.6 |
| 13 | EGK.VI_4 | 111.2 | 50 | 5560.2 | 0.6 | 7.4 | 6 | 926.7 |
| 14 | EGK.VI_5 | 141.7 | 50 | 7083.5 | 0.7 | 6.6 | 6 | 1180.6 |
| 15 | EGK.VI_6 | 144.9 | 50 | 7245.4 | 0.7 | 6.5 | 7 | 1035.1 |
| 16 | EGK.VIII_4 | 310.2 | 30 | 9304.9 | 0.9 | 7.6 | 5 | 1861 |
| 17 | EGK.VIII_5 | 165.2 | 30 | 4957.5 | 0.7 | 7.2 | 5 | 991.5 |
| 18 | EGK.VIII_7 | 63.8 | 25 | 1594.8 | 0.2 | 6.8 | 3 | 531.6 |
| 19 | EGK.X_1 | 386.1 | 50 | 19306.9 | 1.4 | 7.9 | 10 | 1930.7 |
| 20 | EGK.X_3 | 459 | 50 | 22948.1 | 1.7 | 7.9 | 10 | 2294.8 |
| 21 | EGK.X_4 | 264.7 | 50 | 13232.6 | 1.1 | 7.9 | 10 | 1323.3 |

**Table 3-2. The number of clean reads from Trimmomatic**

| Sample ID | # of surviving reads | Surviving rate |
|---|---|---|
| Oocyte_1 | 56,024,575 | 94.81% |
| Oocyte_2 | 56,043,780 | 94.14% |
| Oocyte_3 | 59,498,675 | 95.54% |
| Zygote_2 | 53,378,148 | 96.74% |
| Zygote_3 | 53,999,584 | 96.77% |
| Zygote_5 | 50,027,929 | 98.02% |
| EGK.I_2 | 56,909,314 | 97.36% |
| EGK.I_4 | 61,447,014 | 97.94% |
| EGK.I_6 | 50,188,847 | 96.80% |
| EGK.III_2 | 60,876,681 | 97.30% |
| EGK.III_5 | 56,357,690 | 97.90% |
| EGK.III_6 | 45,715,485 | 98.02% |
| EGK.VI_4 | 62,075,038 | 97.53% |
| EGK.VI_5 | 65,223,164 | 97.77% |
| EGK.VI_6 | 49,604,292 | 98.16% |
| EGK.VIII_4 | 67,401,388 | 97.35% |
| EGK.VIII_5 | 56,396,268 | 96.82% |
| EGK.VIII_7 | 71,309,063 | 97.44% |
| EGK.X_1 | 67,742,313 | 95.72% |
| EGK.X_3 | 74,117,929 | 95.03% |
| EGK.X_4 | 63,233,496 | 94.66% |

**Table 3-3. The number of mapped reads from Tophat2**

| Sample ID | # of mapped reads | Mapping rates |
|---|---|---|
| Oocyte_1 | 38,712,981 | 69.10% |
| Oocyte_2 | 38,950,427 | 69.50% |
| Oocyte_3 | 40,756,592 | 68.50% |
| Zygote_2 | 36,670,787 | 68.70% |
| Zygote_3 | 37,313,712 | 69.10% |
| Zygote_5 | 33,418,656 | 66.80% |
| EGK.I_2 | 33,291,948 | 58.50% |
| EGK.I_4 | 35,024,797 | 57.00% |
| EGK.I_6 | 34,028,038 | 67.80% |
| EGK.III_2 | 37,013,022 | 60.80% |
| EGK.III_5 | 33,532,825 | 59.50% |
| EGK.III_6 | 27,292,144 | 59.70% |
| EGK.VI_4 | 33,148,070 | 53.40% |
| EGK.VI_5 | 44,156,082 | 67.70% |
| EGK.VI_6 | 29,564,158 | 59.60% |
| EGK.VIII_4 | 34,913,918 | 51.80% |
| EGK.VIII_5 | 26,901,019 | 47.70% |
| EGK.VIII_7 | 37,437,258 | 52.50% |
| EGK.X_1 | 36,784,075 | 54.30% |
| EGK.X_3 | 38,689,558 | 52.20% |
| EGK.X_4 | 33,513,752 | 53.00% |

**Table 3-4. Significantly detected BP of GO and KEGG based on up- and down-regulated DEGs between oocyte and zygote**

**Significantly detected BP terms of GO based on up-regulated DEGs between oocyte and zygote (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| DNA metabolic process | 107 | 2.5 | 3.20E-13 |
| cell cycle | 77 | 1.8 | 1.60E-11 |
| protein amino acid phosphorylation | 156 | 3.6 | 2.50E-11 |
| response to DNA damage stimulus | 69 | 1.6 | 4.30E-11 |
| DNA replication | 54 | 1.2 | 4.40E-11 |
| DNA repair | 58 | 1.3 | 2.20E-10 |
| chromosome organization | 73 | 1.7 | 4.00E-09 |
| phosphate metabolic process | 193 | 4.5 | 9.30E-09 |
| phosphorus metabolic process | 193 | 4.5 | 9.30E-09 |
| phosphorylation | 167 | 3.9 | 1.30E-08 |
| cellular response to stress | 76 | 1.8 | 1.90E-08 |
| M phase | 35 | 0.8 | 6.20E-08 |
| microtubule cytoskeleton organization | 26 | 0.6 | 1.10E-07 |
| cell cycle phase | 42 | 1 | 1.40E-07 |
| cell cycle process | 53 | 1.2 | 1.90E-07 |
| microtubule-based process | 47 | 1.1 | 7.20E-07 |
| DNA-dependent DNA replication | 20 | 0.5 | 1.50E-06 |
| regulation of cell cycle | 41 | 0.9 | 0.000004 |
| regulation of small GTPase mediated signal transduction | 57 | 1.3 | 6.70E-06 |
| chromatin modification | 37 | 0.9 | 0.000013 |
| chromatin organization | 53 | 1.2 | 0.000015 |
| mitotic cell cycle | 31 | 0.7 | 0.00002 |
| regulation of DNA metabolic process | 23 | 0.5 | 0.000027 |
| M phase of mitotic cell cycle | 23 | 0.5 | 0.000027 |
| intracellular signaling cascade | 133 | 3.1 | 0.000032 |
| cell division | 29 | 0.7 | 0.000056 |
| regulation of transcription | 268 | 6.2 | 0.000059 |
| determination of bilateral symmetry | 15 | 0.3 | 0.000097 |
| determination of symmetry | 15 | 0.3 | 0.000097 |
| transcription | 125 | 2.9 | 0.00011 |
| nuclear division | 21 | 0.5 | 0.00014 |
| mitosis | 21 | 0.5 | 0.00014 |
| organelle fission | 22 | 0.5 | 0.00017 |
| determination of left/right symmetry | 14 | 0.3 | 0.00026 |
| regulation of Ras protein signal transduction | 44 | 1 | 0.00028 |
| positive regulation of cell cycle | 15 | 0.3 | 0.0004 |
| cytoskeleton organization | 49 | 1.1 | 0.00042 |
| RNA processing | 55 | 1.3 | 0.00045 |
| covalent chromatin modification | 21 | 0.5 | 0.00057 |
| double-strand break repair | 17 | 0.4 | 0.00072 |
| histone modification | 20 | 0.5 | 0.00076 |
| regulation of GTPase activity | 25 | 0.6 | 0.00084 |

| | | | |
|---|---|---|---|
| positive regulation of macromolecule metabolic process | 95 | 2.2 | 0.00088 |
| regulation of Ras GTPase activity | 23 | 0.5 | 0.001 |
| regulation of DNA replication | 13 | 0.3 | 0.0013 |
| negative regulation of programmed cell death | 43 | 1 | 0.0015 |
| negative regulation of cell death | 43 | 1 | 0.0015 |
| chordate embryonic development | 58 | 1.3 | 0.0015 |
| embryonic development ending in birth or egg hatching | 58 | 1.3 | 0.0015 |
| epithelium development | 33 | 0.8 | 0.0018 |
| germ cell development | 17 | 0.4 | 0.0018 |
| negative regulation of apoptosis | 42 | 1 | 0.0019 |
| macromolecular complex subunit organization | 55 | 1.3 | 0.0021 |
| morphogenesis of an epithelium | 23 | 0.5 | 0.0021 |
| cell proliferation | 34 | 0.8 | 0.0023 |
| mRNA metabolic process | 29 | 0.7 | 0.0023 |
| cellular macromolecular complex subunit organization | 39 | 0.9 | 0.0025 |
| anti-apoptosis | 21 | 0.5 | 0.0026 |
| DNA recombination | 17 | 0.4 | 0.0028 |
| positive regulation of nitrogen compound metabolic process | 77 | 1.8 | 0.0028 |
| protein kinase cascade | 31 | 0.7 | 0.003 |
| ubiquitin-dependent protein catabolic process | 31 | 0.7 | 0.003 |
| cell projection organization | 40 | 0.9 | 0.003 |
| gamete generation | 29 | 0.7 | 0.003 |
| regulation of mitotic cell cycle | 18 | 0.4 | 0.0031 |
| microtubule organizing center organization | 9 | 0.2 | 0.0035 |
| regulation of translation | 16 | 0.4 | 0.0038 |
| protein amino acid acylation | 11 | 0.3 | 0.004 |
| macromolecular complex assembly | 51 | 1.2 | 0.0041 |
| positive regulation of macromolecule biosynthetic process | 77 | 1.8 | 0.0042 |
| sexual reproduction | 33 | 0.8 | 0.0046 |
| positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 74 | 1.7 | 0.0049 |
| negative regulation of DNA replication | 7 | 0.2 | 0.005 |
| histone H3 acetylation | 7 | 0.2 | 0.005 |
| histone acetylation | 10 | 0.2 | 0.0053 |
| protein amino acid acetylation | 10 | 0.2 | 0.0053 |
| DNA replication initiation | 6 | 0.1 | 0.0054 |
| in utero embryonic development | 31 | 0.7 | 0.006 |
| cellular macromolecular complex assembly | 35 | 0.8 | 0.0066 |
| posttranscriptional regulation of gene expression | 20 | 0.5 | 0.0067 |
| tissue morphogenesis | 33 | 0.8 | 0.007 |
| negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 49 | 1.1 | 0.0073 |
| modification-dependent macromolecule catabolic process | 44 | 1 | 0.0078 |
| modification-dependent protein catabolic process | 44 | 1 | 0.0078 |
| spermatogenesis | 19 | 0.4 | 0.0089 |
| male gamete generation | 19 | 0.4 | 0.0089 |
| DNA damage checkpoint | 8 | 0.2 | 0.0091 |
| DNA integrity checkpoint | 8 | 0.2 | 0.0091 |
| positive regulation of cellular biosynthetic process | 77 | 1.8 | 0.0098 |
| negative regulation of nitrogen compound metabolic process | 49 | 1.1 | 0.0099 |

| | | | |
|---|---|---|---|
| regulation of gene expression, epigenetic | 14 | 0.3 | 0.01 |
| neuron differentiation | 47 | 1.1 | 0.011 |
| positive regulation of biosynthetic process | 77 | 1.8 | 0.011 |
| DNA damage response, signal transduction | 11 | 0.3 | 0.011 |
| regulation of transcription from RNA polymerase II promoter | 69 | 1.6 | 0.012 |
| regulation of gene-specific transcription | 17 | 0.4 | 0.012 |
| protein modification by small protein conjugation or removal | 18 | 0.4 | 0.012 |
| sensory organ development | 38 | 0.9 | 0.012 |
| cell cycle checkpoint | 9 | 0.2 | 0.013 |
| negative regulation of DNA metabolic process | 9 | 0.2 | 0.013 |
| regulation of cell proliferation | 76 | 1.8 | 0.013 |
| cell projection assembly | 14 | 0.3 | 0.015 |
| chromosome segregation | 10 | 0.2 | 0.015 |
| neural tube development | 16 | 0.4 | 0.016 |
| regulation of apoptosis | 69 | 1.6 | 0.016 |
| protein amino acid deacetylation | 8 | 0.2 | 0.017 |
| meiotic cell cycle | 8 | 0.2 | 0.017 |
| spermatid development | 8 | 0.2 | 0.017 |
| M phase of meiotic cell cycle | 8 | 0.2 | 0.017 |
| meiosis | 8 | 0.2 | 0.017 |
| regulation of programmed cell death | 70 | 1.6 | 0.017 |
| microtubule-based movement | 22 | 0.5 | 0.019 |
| regulation of cell death | 70 | 1.6 | 0.019 |
| regulation of RNA metabolic process | 188 | 4.4 | 0.02 |
| positive regulation of organelle organization | 13 | 0.3 | 0.02 |
| immune effector process | 13 | 0.3 | 0.02 |
| urogenital system development | 20 | 0.5 | 0.02 |
| RNA splicing | 18 | 0.4 | 0.021 |
| telomere organization | 7 | 0.2 | 0.023 |
| central nervous system neuron development | 7 | 0.2 | 0.023 |
| telomere maintenance | 7 | 0.2 | 0.023 |
| double-strand break repair via homologous recombination | 7 | 0.2 | 0.023 |
| centrosome organization | 7 | 0.2 | 0.023 |
| recombinational repair | 7 | 0.2 | 0.023 |
| neuron projection development | 27 | 0.6 | 0.024 |
| regulation of MAPKKK cascade | 10 | 0.2 | 0.024 |
| reproductive process in a multicellular organism | 34 | 0.8 | 0.024 |
| multicellular organism reproduction | 34 | 0.8 | 0.024 |
| protein complex biogenesis | 34 | 0.8 | 0.024 |
| protein complex assembly | 34 | 0.8 | 0.024 |
| positive regulation of gene expression | 65 | 1.5 | 0.025 |
| regulation of cellular response to stress | 11 | 0.3 | 0.026 |
| negative regulation of macromolecule biosynthetic process | 49 | 1.1 | 0.026 |
| regulation of transcription, DNA-dependent | 184 | 4.3 | 0.026 |
| kidney development | 18 | 0.4 | 0.027 |
| positive regulation of specific transcription from RNA polymerase II promoter | 12 | 0.3 | 0.027 |
| regulation of Rab protein signal transduction | 14 | 0.3 | 0.028 |
| regulation of cell cycle process | 14 | 0.3 | 0.028 |
| regulation of Rab GTPase activity | 14 | 0.3 | 0.028 |

| | | | |
|---|---|---|---|
| negative regulation of biosynthetic process | 51 | 1.2 | 0.03 |
| regulation of organelle organization | 26 | 0.6 | 0.03 |
| chromatin assembly or disassembly | 21 | 0.5 | 0.03 |
| reproductive cellular process | 21 | 0.5 | 0.03 |
| negative regulation of cellular biosynthetic process | 50 | 1.2 | 0.031 |
| regulation of cyclin-dependent protein kinase activity | 6 | 0.1 | 0.031 |
| somatic diversification of immunoglobulins | 6 | 0.1 | 0.031 |
| somatic recombination of immunoglobulin gene segments | 6 | 0.1 | 0.031 |
| cellular component morphogenesis | 40 | 0.9 | 0.032 |
| leukocyte mediated immunity | 9 | 0.2 | 0.033 |
| embryonic morphogenesis | 49 | 1.1 | 0.033 |
| regulation of hydrolase activity | 33 | 0.8 | 0.035 |
| proteolysis involved in cellular protein catabolic process | 45 | 1 | 0.036 |
| positive regulation of DNA metabolic process | 11 | 0.3 | 0.036 |
| response to UV | 11 | 0.3 | 0.036 |
| morphogenesis of embryonic epithelium | 12 | 0.3 | 0.037 |
| positive regulation of gene-specific transcription | 13 | 0.3 | 0.037 |
| positive regulation of transcription | 62 | 1.4 | 0.04 |
| positive regulation of MAPKKK cascade | 7 | 0.2 | 0.04 |
| negative regulation of cell proliferation | 30 | 0.7 | 0.04 |
| regulation of cellular protein metabolic process | 40 | 0.9 | 0.041 |
| cellular protein complex assembly | 18 | 0.4 | 0.042 |
| protein modification by small protein removal | 5 | 0.1 | 0.042 |
| postreplication repair | 5 | 0.1 | 0.042 |
| negative regulation of RNA metabolic process | 32 | 0.7 | 0.042 |
| negative regulation of transcription, DNA-dependent | 31 | 0.7 | 0.044 |
| vasculogenesis | 8 | 0.2 | 0.045 |
| spermatid differentiation | 8 | 0.2 | 0.045 |
| pancreas development | 8 | 0.2 | 0.045 |
| central nervous system neuron differentiation | 8 | 0.2 | 0.045 |
| mRNA processing | 21 | 0.5 | 0.045 |
| cellular protein catabolic process | 45 | 1 | 0.046 |
| negative regulation of macromolecule metabolic process | 59 | 1.4 | 0.046 |
| response to ionizing radiation | 9 | 0.2 | 0.048 |
| regulation of specific transcription from RNA polymerase II promoter | 13 | 0.3 | 0.049 |
| cellular macromolecule catabolic process | 54 | 1.2 | 0.049 |
| regulation of ARF protein signal transduction | 10 | 0.2 | 0.049 |
| gene silencing | 10 | 0.2 | 0.049 |
| dephosphorylation | 29 | 0.7 | 0.049 |
| regulation of DNA binding | 12 | 0.3 | 0.05 |
| cell part morphogenesis | 26 | 0.6 | 0.05 |

**Significantly detected KEGG pathways based on up-regulated DEGs between oocyte and zygote (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Cell cycle | 55 | 1.3 | 2.90E-08 |
| DNA replication | 19 | 0.4 | 5.10E-06 |
| Notch signaling pathway | 24 | 0.6 | 1.20E-04 |
| Spliceosome | 45 | 1 | 1.30E-04 |
| Homologous recombination | 15 | 0.3 | 4.60E-04 |

| | | | |
|---|---|---|---|
| Progesterone-mediated oocyte maturation | 32 | 0.7 | 8.50E-04 |
| MAPK signaling pathway | 76 | 1.8 | 1.30E-03 |
| Base excision repair | 14 | 0.3 | 2.00E-03 |
| Dorso-ventral axis formation | 12 | 0.3 | 3.10E-03 |
| Ubiquitin mediated proteolysis | 45 | 1 | 3.30E-03 |
| Oocyte meiosis | 37 | 0.9 | 3.90E-03 |
| ErbB signaling pathway | 32 | 0.7 | 4.50E-03 |
| Regulation of actin cytoskeleton | 60 | 1.4 | 7.50E-03 |
| Mismatch repair | 10 | 0.2 | 8.80E-03 |
| Toll-like receptor signaling pathway | 31 | 0.7 | 9.10E-03 |
| Insulin signaling pathway | 41 | 0.9 | 1.30E-02 |
| Basal transcription factors | 13 | 0.3 | 1.70E-02 |
| Pyrimidine metabolism | 31 | 0.7 | 1.70E-02 |
| Nucleotide excision repair | 15 | 0.3 | 2.10E-02 |
| Non-homologous end-joining | 8 | 0.2 | 2.50E-02 |
| Wnt signaling pathway | 44 | 1 | 3.10E-02 |
| Glutathione metabolism | 15 | 0.3 | 3.80E-02 |
| p53 signaling pathway | 22 | 0.5 | 4.00E-02 |
| Endocytosis | 54 | 1.2 | 4.40E-02 |
| VEGF signaling pathway | 23 | 0.5 | 4.90E-02 |

**Significantly detected BP terms of GO based on down-regulated DEGs between oocyte and zygote (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| transmembrane transport | 116 | 3.5 | 4.20E-10 |
| ion transport | 116 | 3.5 | 9.50E-06 |
| metal ion transport | 68 | 2 | 1.40E-05 |
| cation transport | 80 | 2.4 | 1.20E-04 |
| oxidation reduction | 100 | 3 | 2.30E-04 |
| sterol metabolic process | 13 | 0.4 | 2.60E-04 |
| response to oxidative stress | 16 | 0.5 | 7.80E-04 |
| response to oxygen levels | 13 | 0.4 | 8.00E-04 |
| amino acid transport | 14 | 0.4 | 8.20E-04 |
| protein folding | 28 | 0.8 | 9.20E-04 |
| homeostatic process | 67 | 2 | 1.10E-03 |
| di-, tri-valent inorganic cation transport | 25 | 0.7 | 1.10E-03 |
| calcium ion transport | 23 | 0.7 | 1.40E-03 |
| cellular homeostasis | 43 | 1.3 | 1.60E-03 |
| cholesterol metabolic process | 11 | 0.3 | 2.00E-03 |
| response to hypoxia | 12 | 0.4 | 2.10E-03 |
| regulation of cell size | 19 | 0.6 | 2.40E-03 |
| amine transport | 15 | 0.4 | 2.60E-03 |
| organic acid transport | 16 | 0.5 | 3.40E-03 |
| carboxylic acid transport | 16 | 0.5 | 3.40E-03 |
| cellular calcium ion homeostasis | 15 | 0.4 | 3.70E-03 |
| translation | 53 | 1.6 | 3.80E-03 |
| regulation of cell growth | 18 | 0.5 | 4.70E-03 |
| calcium ion homeostasis | 15 | 0.4 | 5.20E-03 |
| lipid biosynthetic process | 29 | 0.9 | 6.40E-03 |
| fatty acid metabolic process | 19 | 0.6 | 6.80E-03 |

| | | | |
|---|---|---|---|
| cellular cation homeostasis | 19 | 0.6 | 6.80E-03 |
| protein maturation | 12 | 0.4 | 7.10E-03 |
| steroid metabolic process | 20 | 0.6 | 7.50E-03 |
| regulation of heart rate | 7 | 0.2 | 7.50E-03 |
| positive regulation of apoptosis | 29 | 0.9 | 7.60E-03 |
| positive regulation of programmed cell death | 29 | 0.9 | 7.60E-03 |
| zinc ion transport | 5 | 0.1 | 7.60E-03 |
| steroid biosynthetic process | 11 | 0.3 | 7.70E-03 |
| cellular di-, tri-valent inorganic cation homeostasis | 18 | 0.5 | 7.90E-03 |
| sodium ion transport | 18 | 0.5 | 7.90E-03 |
| chemical homeostasis | 39 | 1.2 | 8.20E-03 |
| death | 36 | 1.1 | 8.20E-03 |
| di-, tri-valent inorganic cation homeostasis | 19 | 0.6 | 8.70E-03 |
| response to wounding | 28 | 0.8 | 8.80E-03 |
| positive regulation of cell death | 29 | 0.9 | 9.00E-03 |
| cellular metal ion homeostasis | 15 | 0.4 | 9.30E-03 |
| regulation of cellular component size | 24 | 0.7 | 9.40E-03 |
| extracellular matrix organization | 16 | 0.5 | 1.10E-02 |
| apoptosis | 32 | 1 | 1.10E-02 |
| cell death | 35 | 1 | 1.10E-02 |
| regulation of blood pressure | 11 | 0.3 | 1.10E-02 |
| antigen processing and presentation of peptide antigen | 6 | 0.2 | 1.20E-02 |
| multicellular organismal metabolic process | 6 | 0.2 | 1.20E-02 |
| sterol biosynthetic process | 6 | 0.2 | 1.20E-02 |
| response to organic substance | 40 | 1.2 | 1.20E-02 |
| regulation of growth | 30 | 0.9 | 1.50E-02 |
| metal ion homeostasis | 15 | 0.4 | 1.60E-02 |
| programmed cell death | 32 | 1 | 1.70E-02 |
| circulatory system process | 17 | 0.5 | 1.80E-02 |
| induction of apoptosis | 17 | 0.5 | 1.80E-02 |
| blood circulation | 17 | 0.5 | 1.80E-02 |
| induction of programmed cell death | 17 | 0.5 | 1.80E-02 |
| cation homeostasis | 22 | 0.7 | 1.80E-02 |
| antigen processing and presentation | 12 | 0.4 | 1.90E-02 |
| regulation of cell morphogenesis | 12 | 0.4 | 1.90E-02 |
| collagen metabolic process | 5 | 0.1 | 1.90E-02 |
| multicellular organismal macromolecule metabolic process | 5 | 0.1 | 1.90E-02 |
| antigen processing and presentation of exogenous peptide antigen | 5 | 0.1 | 1.90E-02 |
| antigen processing and presentation of exogenous antigen | 5 | 0.1 | 1.90E-02 |
| enzyme linked receptor protein signaling pathway | 39 | 1.2 | 2.00E-02 |
| negative regulation of immune system process | 8 | 0.2 | 2.10E-02 |
| aminoglycan metabolic process | 8 | 0.2 | 2.10E-02 |
| polyol metabolic process | 8 | 0.2 | 2.10E-02 |
| pigmentation during development | 6 | 0.2 | 2.20E-02 |
| phospholipid transport | 7 | 0.2 | 2.30E-02 |
| carbohydrate catabolic process | 17 | 0.5 | 2.30E-02 |
| cell redox homeostasis | 14 | 0.4 | 2.30E-02 |
| protein processing | 10 | 0.3 | 2.50E-02 |
| transmembrane receptor protein serine/threonine kinase | 15 | 0.4 | 2.50E-02 |

| | | | |
|---|---|---|---|
| signaling pathway | | | |
| monovalent inorganic cation transport | 45 | 1.3 | 2.70E-02 |
| response to axon injury | 4 | 0.1 | 3.00E-02 |
| cellular ion homeostasis | 26 | 0.8 | 3.00E-02 |
| cellular chemical homeostasis | 26 | 0.8 | 3.00E-02 |
| homeostasis of number of cells | 15 | 0.4 | 3.10E-02 |
| regulation of angiogenesis | 8 | 0.2 | 3.10E-02 |
| lipid localization | 16 | 0.5 | 3.20E-02 |
| T cell activation | 16 | 0.5 | 3.20E-02 |
| response to endogenous stimulus | 18 | 0.5 | 3.40E-02 |
| regulation of binding | 13 | 0.4 | 3.40E-02 |
| cell growth | 7 | 0.2 | 3.50E-02 |
| pigmentation | 7 | 0.2 | 3.50E-02 |
| lymphocyte activation | 21 | 0.6 | 3.50E-02 |
| proteolysis | 96 | 2.9 | 3.60E-02 |
| associative learning | 5 | 0.1 | 3.70E-02 |
| pigment cell differentiation | 5 | 0.1 | 3.70E-02 |
| translational elongation | 6 | 0.2 | 3.70E-02 |
| oxygen and reactive oxygen species metabolic process | 6 | 0.2 | 3.70E-02 |
| regulation of cell adhesion | 15 | 0.4 | 3.80E-02 |
| transforming growth factor beta receptor signaling pathway | 9 | 0.3 | 3.80E-02 |
| regulation of apoptosis | 55 | 1.6 | 3.90E-02 |
| leukocyte activation | 23 | 0.7 | 4.10E-02 |
| potassium ion transport | 26 | 0.8 | 4.50E-02 |
| fatty acid biosynthetic process | 11 | 0.3 | 4.80E-02 |
| neurotransmitter transport | 11 | 0.3 | 4.80E-02 |

**Significantly detected KEGG pathways based on down-regulated DEGs between oocyte and zygote (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Ribosome | 48 | 1.4 | 1.20E-15 |
| Lysosome | 46 | 1.4 | 1.30E-08 |
| Glycosaminoglycan degradation | 12 | 0.4 | 7.30E-04 |
| Steroid biosynthesis | 9 | 0.3 | 1.90E-03 |
| Other glycan degradation | 8 | 0.2 | 1.00E-02 |
| Glycosphingolipid biosynthesis | 8 | 0.2 | 1.00E-02 |
| Starch and sucrose metabolism | 11 | 0.3 | 2.80E-02 |
| N-Glycan biosynthesis | 15 | 0.4 | 2.90E-02 |
| ECM-receptor interaction | 24 | 0.7 | 3.60E-02 |

**Table 3-5. Significantly detected BP of GO and KEGG based on up-regulated DEGs between EGK.III and EGK.VI**

**Significantly detected BP terms of GO based on up-regulated DEGs between EGK.III and EGK.VI (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| embryonic morphogenesis | 29 | 3 | 3.60E-08 |
| pattern specification process | 26 | 2.7 | 1.60E-07 |
| neuron differentiation | 24 | 2.5 | 3.10E-06 |
| embryonic organ morphogenesis | 17 | 1.8 | 4.50E-06 |
| regionalization | 20 | 2.1 | 5.10E-06 |
| forebrain development | 16 | 1.7 | 5.50E-06 |
| embryonic organ development | 19 | 2 | 1.10E-05 |
| anterior/posterior pattern formation | 16 | 1.7 | 2.10E-05 |
| neuron development | 18 | 1.9 | 4.10E-05 |
| regulation of RNA metabolic process | 63 | 6.6 | 9.20E-05 |
| positive regulation of macromolecule metabolic process | 34 | 3.6 | 1.10E-04 |
| axonogenesis | 13 | 1.4 | 1.50E-04 |
| enzyme linked receptor protein signaling pathway | 21 | 2.2 | 1.50E-04 |
| chordate embryonic development | 23 | 2.4 | 1.50E-04 |
| embryonic development ending in birth or egg hatching | 23 | 2.4 | 1.50E-04 |
| metal ion transport | 27 | 2.8 | 1.90E-04 |
| regulation of transcription, DNA-dependent | 61 | 6.4 | 2.00E-04 |
| potassium ion transport | 16 | 1.7 | 2.00E-04 |
| neuron projection morphogenesis | 13 | 1.4 | 2.10E-04 |
| cell-cell signaling | 17 | 1.8 | 2.10E-04 |
| neuron projection development | 14 | 1.5 | 2.50E-04 |
| positive regulation of gene expression | 26 | 2.7 | 3.90E-04 |
| cell morphogenesis involved in neuron differentiation | 13 | 1.4 | 4.00E-04 |
| cell projection morphogenesis | 13 | 1.4 | 5.30E-04 |
| tube development | 17 | 1.8 | 5.40E-04 |
| cell morphogenesis involved in differentiation | 14 | 1.5 | 5.80E-04 |
| positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 27 | 2.8 | 5.80E-04 |
| positive regulation of transcription | 25 | 2.6 | 5.90E-04 |
| Wnt receptor signaling pathway | 12 | 1.3 | 8.40E-04 |
| positive regulation of nitrogen compound metabolic process | 27 | 2.8 | 8.80E-04 |
| gland development | 12 | 1.3 | 9.70E-04 |
| cell part morphogenesis | 13 | 1.4 | 1.00E-03 |
| positive regulation of macromolecule biosynthetic process | 27 | 2.8 | 1.10E-03 |
| transmembrane receptor protein tyrosine kinase signaling pathway | 15 | 1.6 | 1.10E-03 |
| ion transport | 40 | 4.2 | 1.10E-03 |
| neurotransmitter transport | 8 | 0.8 | 1.10E-03 |
| inner ear morphogenesis | 9 | 0.9 | 1.20E-03 |
| sensory organ development | 16 | 1.7 | 1.20E-03 |
| Wnt receptor signaling pathway, calcium modulating pathway | 7 | 0.7 | 1.20E-03 |
| ear morphogenesis | 9 | 0.9 | 1.40E-03 |
| positive regulation of RNA metabolic process | 21 | 2.2 | 1.60E-03 |

| | | | |
|---|---|---|---|
| positive regulation of cellular biosynthetic process | 27 | 2.8 | 1.70E-03 |
| regulation of transcription | 73 | 7.6 | 1.70E-03 |
| positive regulation of biosynthetic process | 27 | 2.8 | 1.80E-03 |
| monovalent inorganic cation transport | 21 | 2.2 | 1.80E-03 |
| axon guidance | 9 | 0.9 | 2.00E-03 |
| appendage development | 11 | 1.1 | 2.00E-03 |
| limb development | 11 | 1.1 | 2.00E-03 |
| ear development | 10 | 1 | 2.10E-03 |
| negative regulation of transcription from RNA polymerase II promoter | 11 | 1.1 | 2.70E-03 |
| response to organic substance | 18 | 1.9 | 2.90E-03 |
| positive regulation of transcription, DNA-dependent | 20 | 2.1 | 3.20E-03 |
| segmentation | 7 | 0.7 | 3.20E-03 |
| regulation of transcription from RNA polymerase II promoter | 24 | 2.5 | 3.20E-03 |
| negative regulation of nitrogen compound metabolic process | 18 | 1.9 | 3.70E-03 |
| cellular component morphogenesis | 16 | 1.7 | 3.90E-03 |
| inner ear development | 9 | 0.9 | 4.50E-03 |
| pituitary gland development | 5 | 0.5 | 4.50E-03 |
| tube morphogenesis | 11 | 1.1 | 4.90E-03 |
| positive regulation of transcription from RNA polymerase II promoter | 18 | 1.9 | 5.00E-03 |
| negative regulation of gene expression | 17 | 1.8 | 5.60E-03 |
| diencephalon development | 5 | 0.5 | 6.20E-03 |
| skeletal system development | 16 | 1.7 | 6.30E-03 |
| cell morphogenesis | 14 | 1.5 | 6.60E-03 |
| somitogenesis | 6 | 0.6 | 7.40E-03 |
| synaptic transmission | 9 | 0.9 | 7.70E-03 |
| cell projection organization | 14 | 1.5 | 7.80E-03 |
| negative regulation of macromolecule metabolic process | 21 | 2.2 | 7.90E-03 |
| regulation of Wnt receptor signaling pathway | 5 | 0.5 | 8.30E-03 |
| cation transport | 27 | 2.8 | 8.30E-03 |
| endocrine system development | 7 | 0.7 | 8.30E-03 |
| negative regulation of cellular biosynthetic process | 18 | 1.9 | 8.70E-03 |
| neurological system process | 19 | 2 | 1.00E-02 |
| negative regulation of biosynthetic process | 18 | 1.9 | 1.10E-02 |
| respiratory system development | 9 | 0.9 | 1.10E-02 |
| negative regulation of transcription | 15 | 1.6 | 1.10E-02 |
| sensory perception | 12 | 1.3 | 1.20E-02 |
| gastrulation | 7 | 0.7 | 1.30E-02 |
| negative regulation of macromolecule biosynthetic process | 17 | 1.8 | 1.40E-02 |
| morphogenesis of a branching structure | 8 | 0.8 | 1.40E-02 |
| neural tube development | 7 | 0.7 | 1.50E-02 |
| appendage morphogenesis | 9 | 0.9 | 1.50E-02 |
| limb morphogenesis | 9 | 0.9 | 1.50E-02 |
| negative regulation of transcription, DNA-dependent | 12 | 1.3 | 1.60E-02 |
| muscle organ development | 10 | 1 | 1.60E-02 |
| blood vessel morphogenesis | 11 | 1.1 | 1.60E-02 |
| negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 16 | 1.7 | 1.70E-02 |
| skeletal system morphogenesis | 9 | 0.9 | 1.70E-02 |
| embryonic limb morphogenesis | 8 | 0.8 | 1.80E-02 |

| | | | |
|---|---|---|---|
| embryonic appendage morphogenesis | 8 | 0.8 | 1.80E-02 |
| negative regulation of cell differentiation | 12 | 1.3 | 1.80E-02 |
| cell motion | 16 | 1.7 | 1.90E-02 |
| hippocampus development | 4 | 0.4 | 1.90E-02 |
| negative regulation of RNA metabolic process | 12 | 1.3 | 2.00E-02 |
| cell differentiation in hindbrain | 3 | 0.3 | 2.00E-02 |
| reproductive developmental process | 12 | 1.3 | 2.10E-02 |
| regulation of cell proliferation | 23 | 2.4 | 2.30E-02 |
| reproductive process in a multicellular organism | 12 | 1.3 | 2.50E-02 |
| multicellular organism reproduction | 12 | 1.3 | 2.50E-02 |
| telencephalon development | 6 | 0.6 | 2.50E-02 |
| adult behavior | 7 | 0.7 | 2.60E-02 |
| transmission of nerve impulse | 9 | 0.9 | 2.80E-02 |
| cell surface receptor linked signal transduction | 61 | 6.4 | 2.80E-02 |
| vasculature development | 12 | 1.3 | 3.10E-02 |
| cognition | 13 | 1.4 | 3.20E-02 |
| positive regulation of cell proliferation | 15 | 1.6 | 3.40E-02 |
| pallium development | 5 | 0.5 | 3.50E-02 |
| branching morphogenesis of a tube | 7 | 0.7 | 3.60E-02 |
| positive regulation of developmental process | 12 | 1.3 | 3.70E-02 |
| regulation of ossification | 6 | 0.6 | 3.70E-02 |
| cell migration | 12 | 1.3 | 4.00E-02 |
| limbic system development | 4 | 0.4 | 4.10E-02 |
| cell growth | 4 | 0.4 | 4.10E-02 |
| regulation of neurotransmitter levels | 4 | 0.4 | 4.10E-02 |
| response to hypoxia | 5 | 0.5 | 4.10E-02 |
| amine transport | 6 | 0.6 | 4.20E-02 |
| response to oxygen levels | 5 | 0.5 | 4.80E-02 |
| regulation of osteoblast differentiation | 5 | 0.5 | 4.80E-02 |
| digestive system development | 4 | 0.4 | 4.90E-02 |
| genitalia development | 4 | 0.4 | 4.90E-02 |
| digestive tract morphogenesis | 4 | 0.4 | 4.90E-02 |

**Significantly detected KEGG pathways based on up-regulated DEGs between EGK.III and EGK.VI (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Wnt signaling pathway | 16 | 1.7 | 4.50E-04 |
| Melanogenesis | 12 | 1.3 | 1.20E-03 |
| Neuroactive ligand-receptor interaction | 19 | 2 | 1.30E-02 |
| Hedgehog signaling pathway | 7 | 0.7 | 1.30E-02 |
| TGF-beta signaling pathway | 9 | 0.9 | 1.70E-02 |

**Table 3-6. Significantly detected BP of GO and KEGG based on up-regulated DEGs between EGK.VI and EGK.VIII**

**Significantly detected BP terms of GO based on up-regulated DEGs between EGK.VI and EGK.VIII (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| ribosome biogenesis | 21 | 0.6 | 2.40E-06 |
| ribonucleoprotein complex biogenesis | 24 | 0.7 | 9.10E-06 |
| translation | 60 | 1.8 | 5.90E-05 |
| metal ion transport | 63 | 1.9 | 3.80E-04 |
| rRNA processing | 14 | 0.4 | 5.20E-04 |
| rRNA metabolic process | 14 | 0.4 | 5.20E-04 |
| ncRNA metabolic process | 31 | 0.9 | 1.00E-03 |
| intracellular transport | 57 | 1.7 | 1.80E-03 |
| ncRNA processing | 22 | 0.7 | 2.20E-03 |
| nucleocytoplasmic transport | 16 | 0.5 | 2.50E-03 |
| nuclear transport | 16 | 0.5 | 2.50E-03 |
| heart development | 33 | 1 | 2.50E-03 |
| amine transport | 15 | 0.5 | 2.70E-03 |
| Wnt receptor signaling pathway | 23 | 0.7 | 3.10E-03 |
| amino acid transport | 13 | 0.4 | 3.20E-03 |
| potassium ion transport | 30 | 0.9 | 3.90E-03 |
| protein folding | 26 | 0.8 | 4.80E-03 |
| RNA processing | 43 | 1.3 | 4.80E-03 |
| ion transport | 102 | 3.1 | 5.50E-03 |
| nuclear export | 7 | 0.2 | 7.60E-03 |
| cation transport | 71 | 2.1 | 9.50E-03 |
| organic acid transport | 15 | 0.5 | 9.50E-03 |
| cell redox homeostasis | 15 | 0.5 | 9.50E-03 |
| carboxylic acid transport | 15 | 0.5 | 9.50E-03 |
| tube development | 33 | 1 | 1.10E-02 |
| monovalent inorganic cation transport | 47 | 1.4 | 1.10E-02 |
| positive regulation of macromolecule metabolic process | 73 | 2.2 | 1.40E-02 |
| cellular component morphogenesis | 35 | 1.1 | 1.50E-02 |
| regulation of cytokine production | 18 | 0.5 | 1.60E-02 |
| cellular homeostasis | 39 | 1.2 | 1.60E-02 |
| cellular di-, tri-valent inorganic cation homeostasis | 17 | 0.5 | 1.90E-02 |
| cell-cell signaling | 30 | 0.9 | 2.00E-02 |
| enzyme linked receptor protein signaling pathway | 39 | 1.2 | 2.10E-02 |
| regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains | 8 | 0.2 | 2.20E-02 |
| regulation of adaptive immune response | 8 | 0.2 | 2.20E-02 |
| morphogenesis of embryonic epithelium | 11 | 0.3 | 2.20E-02 |
| pyruvate metabolic process | 6 | 0.2 | 2.30E-02 |
| myofibril assembly | 6 | 0.2 | 2.30E-02 |
| monosaccharide biosynthetic process | 7 | 0.2 | 2.30E-02 |
| tube closure | 7 | 0.2 | 2.30E-02 |
| neural tube closure | 7 | 0.2 | 2.30E-02 |

| | | | |
|---|---|---|---|
| positive regulation of biosynthetic process | 62 | 1.9 | 2.40E-02 |
| positive regulation of macromolecule biosynthetic process | 60 | 1.8 | 2.60E-02 |
| tissue morphogenesis | 26 | 0.8 | 2.70E-02 |
| cellular calcium ion homeostasis | 13 | 0.4 | 2.80E-02 |
| heart morphogenesis | 13 | 0.4 | 2.80E-02 |
| striated muscle cell development | 9 | 0.3 | 2.80E-02 |
| response to wounding | 26 | 0.8 | 3.10E-02 |
| in utero embryonic development | 24 | 0.7 | 3.10E-02 |
| skeletal system development | 35 | 1.1 | 3.10E-02 |
| transition metal ion transport | 8 | 0.2 | 3.20E-02 |
| positive regulation of cellular biosynthetic process | 61 | 1.8 | 3.20E-02 |
| lung development | 16 | 0.5 | 3.30E-02 |
| cellular cation homeostasis | 17 | 0.5 | 3.40E-02 |
| angiogenesis | 17 | 0.5 | 3.40E-02 |
| calcium ion homeostasis | 13 | 0.4 | 3.50E-02 |
| primary neural tube formation | 7 | 0.2 | 3.50E-02 |
| alcohol biosynthetic process | 7 | 0.2 | 3.50E-02 |
| positive regulation of immune system process | 20 | 0.6 | 3.60E-02 |
| forebrain development | 21 | 0.6 | 3.60E-02 |
| hexose biosynthetic process | 6 | 0.2 | 3.80E-02 |
| regulation of lipase activity | 6 | 0.2 | 3.80E-02 |
| translational elongation | 6 | 0.2 | 3.80E-02 |
| regulation of cell adhesion | 15 | 0.5 | 3.90E-02 |
| neuron development | 28 | 0.8 | 4.00E-02 |
| respiratory tube development | 16 | 0.5 | 4.00E-02 |
| sodium ion transport | 16 | 0.5 | 4.00E-02 |
| di-, tri-valent inorganic cation homeostasis | 17 | 0.5 | 4.10E-02 |
| placenta development | 10 | 0.3 | 4.40E-02 |
| negative regulation of nitrogen compound metabolic process | 38 | 1.1 | 4.50E-02 |
| transmembrane receptor protein tyrosine kinase signaling pathway | 27 | 0.8 | 4.60E-02 |
| embryonic morphogenesis | 40 | 1.2 | 4.70E-02 |
| telencephalon development | 11 | 0.3 | 4.80E-02 |
| regulation of cytokine biosynthetic process | 11 | 0.3 | 4.80E-02 |
| respiratory system development | 17 | 0.5 | 4.90E-02 |

**Significantly detected KEGG pathways based on up-regulated DEGs between EGK.VI and EGK.VIII (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Ribosome | 34 | 1 | 4.70E-06 |
| Melanogenesis | 30 | 0.9 | 6.40E-03 |
| Proteasome | 14 | 0.4 | 3.00E-02 |
| Wnt signaling pathway | 38 | 1.1 | 3.70E-02 |
| Calcium signaling pathway | 42 | 1.3 | 4.60E-02 |

**Significantly detected BP terms of GO based on down-regulated DEGs between EGK.VI and EGK.VIII (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| phosphate metabolic process | 120 | 4.4 | 3.60E-05 |
| phosphorus metabolic process | 120 | 4.4 | 3.60E-05 |

| | | | |
|---|---|---|---|
| protein amino acid phosphorylation | 90 | 3.3 | 9.40E-05 |
| regulation of small GTPase mediated signal transduction | 38 | 1.4 | 1.90E-04 |
| regulation of Ras protein signal transduction | 32 | 1.2 | 2.80E-04 |
| intracellular signaling cascade | 86 | 3.1 | 6.10E-04 |
| phosphorylation | 97 | 3.5 | 1.00E-03 |
| regulation of Rab protein signal transduction | 13 | 0.5 | 1.50E-03 |
| regulation of Rab GTPase activity | 13 | 0.5 | 1.50E-03 |
| regulation of Ras GTPase activity | 17 | 0.6 | 1.60E-03 |
| regulation of GTPase activity | 18 | 0.7 | 2.00E-03 |
| dephosphorylation | 24 | 0.9 | 3.80E-03 |
| vesicle-mediated transport | 36 | 1.3 | 5.10E-03 |
| glycoprotein biosynthetic process | 19 | 0.7 | 5.30E-03 |
| reproductive process in a multicellular organism | 26 | 0.9 | 5.80E-03 |
| multicellular organism reproduction | 26 | 0.9 | 5.80E-03 |
| phosphatidylinositol metabolic process | 5 | 0.2 | 7.80E-03 |
| sexual reproduction | 23 | 0.8 | 8.20E-03 |
| glycoprotein metabolic process | 20 | 0.7 | 8.30E-03 |
| regulation of ARF protein signal transduction | 9 | 0.3 | 9.40E-03 |
| glycosylation | 17 | 0.6 | 9.50E-03 |
| biopolymer glycosylation | 17 | 0.6 | 9.50E-03 |
| protein amino acid glycosylation | 17 | 0.6 | 9.50E-03 |
| protein kinase cascade | 21 | 0.8 | 1.00E-02 |
| regulation of phosphate metabolic process | 32 | 1.2 | 1.00E-02 |
| regulation of phosphorus metabolic process | 32 | 1.2 | 1.00E-02 |
| regulation of phosphorylation | 31 | 1.1 | 1.00E-02 |
| positive regulation of cellular component organization | 15 | 0.5 | 1.20E-02 |
| osteoblast differentiation | 8 | 0.3 | 1.30E-02 |
| lipopolysaccharide metabolic process | 4 | 0.1 | 1.50E-02 |
| lipopolysaccharide biosynthetic process | 4 | 0.1 | 1.50E-02 |
| secretion by cell | 16 | 0.6 | 1.50E-02 |
| protein secretion | 5 | 0.2 | 1.60E-02 |
| positive regulation of macromolecule metabolic process | 59 | 2.1 | 1.60E-02 |
| gamete generation | 19 | 0.7 | 1.80E-02 |
| transmembrane receptor protein serine/threonine kinase signaling pathway | 13 | 0.5 | 2.10E-02 |
| secretion | 18 | 0.7 | 2.30E-02 |
| coenzyme metabolic process | 18 | 0.7 | 2.30E-02 |
| transcription from RNA polymerase II promoter | 9 | 0.3 | 2.30E-02 |
| histone acetylation | 7 | 0.3 | 2.40E-02 |
| protein amino acid acetylation | 7 | 0.3 | 2.40E-02 |
| regulation of protein modification process | 19 | 0.7 | 2.70E-02 |
| monosaccharide metabolic process | 22 | 0.8 | 2.80E-02 |
| regulation of interleukin-2 production | 5 | 0.2 | 2.80E-02 |
| histone H3 acetylation | 5 | 0.2 | 2.80E-02 |
| endocytosis | 12 | 0.4 | 2.80E-02 |
| membrane invagination | 12 | 0.4 | 2.80E-02 |
| regulation of protein amino acid phosphorylation | 16 | 0.6 | 3.00E-02 |
| regulation of protein kinase activity | 21 | 0.8 | 3.10E-02 |
| gonad development | 11 | 0.4 | 3.20E-02 |
| female gonad development | 8 | 0.3 | 3.20E-02 |

| | | | |
|---|---|---|---|
| regulation of cellular protein metabolic process | 28 | 1 | 3.20E-02 |
| myeloid leukocyte activation | 6 | 0.2 | 3.20E-02 |
| acetyl-CoA catabolic process | 6 | 0.2 | 3.20E-02 |
| vesicle organization | 6 | 0.2 | 3.20E-02 |
| lung alveolus development | 6 | 0.2 | 3.20E-02 |
| tricarboxylic acid cycle | 6 | 0.2 | 3.20E-02 |
| acetyl-CoA metabolic process | 6 | 0.2 | 3.20E-02 |
| cellular polysaccharide metabolic process | 7 | 0.3 | 3.30E-02 |
| regulation of hydrolase activity | 23 | 0.8 | 3.50E-02 |
| membrane organization | 19 | 0.7 | 3.60E-02 |
| regulation of apoptosis | 45 | 1.6 | 3.60E-02 |
| protein localization | 54 | 2 | 3.80E-02 |
| regulation of kinase activity | 21 | 0.8 | 4.00E-02 |
| regulation of synaptic transmission | 8 | 0.3 | 4.20E-02 |
| regulation of developmental growth | 8 | 0.3 | 4.20E-02 |
| positive regulation of developmental process | 24 | 0.9 | 4.30E-02 |
| cellular polysaccharide biosynthetic process | 5 | 0.2 | 4.40E-02 |
| aromatic compound biosynthetic process | 5 | 0.2 | 4.40E-02 |
| cofactor catabolic process | 7 | 0.3 | 4.50E-02 |
| protein amino acid acylation | 7 | 0.3 | 4.50E-02 |
| negative regulation of molecular function | 15 | 0.5 | 4.50E-02 |
| regulation of synaptic plasticity | 6 | 0.2 | 4.60E-02 |
| negative regulation of apoptosis | 25 | 0.9 | 4.60E-02 |
| establishment of protein localization | 47 | 1.7 | 4.80E-02 |
| protein transport | 47 | 1.7 | 4.80E-02 |
| regulation of programmed cell death | 45 | 1.6 | 4.80E-02 |
| regulation of growth | 23 | 0.8 | 4.80E-02 |
| reproductive structure development | 12 | 0.4 | 4.90E-02 |
| regulation of transferase activity | 21 | 0.8 | 5.00E-02 |

**Significantly detected KEGG pathways based on down-regulated DEGs between EGK.III and EGK.VI (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Endocytosis | 46 | 1.7 | 1.60E-03 |
| Glycerolipid metabolism | 16 | 0.6 | 2.10E-03 |
| Lysine degradation | 13 | 0.5 | 1.20E-02 |
| Valine, leucine and isoleucine degradation | 14 | 0.5 | 1.30E-02 |
| MAPK signaling pathway | 51 | 1.9 | 2.20E-02 |
| Propanoate metabolism | 10 | 0.4 | 2.80E-02 |
| Notch signaling pathway | 14 | 0.5 | 3.80E-02 |

**Table 3-7. Significantly detected BP of GO and KEGG based on up-regulated DEGs between EGK.VIII and EGK.X**

**Significantly detected BP terms of GO based on up-regulated DEGs between EGK.VIII and EGK.X (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| transmembrane transport | 85 | 3.3 | 2.00E-05 |
| tube development | 34 | 1.3 | 1.40E-04 |
| chemical homeostasis | 39 | 1.5 | 1.40E-04 |
| homeostatic process | 59 | 2.3 | 3.00E-04 |
| tube morphogenesis | 23 | 0.9 | 4.40E-04 |
| ion homeostasis | 31 | 1.2 | 4.50E-04 |
| epithelial tube morphogenesis | 14 | 0.5 | 5.50E-04 |
| response to wounding | 27 | 1.1 | 8.80E-04 |
| ion transport | 89 | 3.5 | 1.10E-03 |
| inflammatory response | 14 | 0.5 | 1.10E-03 |
| morphogenesis of an epithelium | 18 | 0.7 | 1.30E-03 |
| tissue morphogenesis | 26 | 1 | 1.70E-03 |
| angiogenesis | 18 | 0.7 | 1.70E-03 |
| primary neural tube formation | 8 | 0.3 | 2.40E-03 |
| epithelium development | 24 | 0.9 | 2.90E-03 |
| cation homeostasis | 21 | 0.8 | 3.60E-03 |
| morphogenesis of embryonic epithelium | 11 | 0.4 | 4.80E-03 |
| lung development | 16 | 0.6 | 4.90E-03 |
| cellular homeostasis | 35 | 1.4 | 5.20E-03 |
| respiratory tube development | 16 | 0.6 | 6.10E-03 |
| transmembrane receptor protein tyrosine kinase signaling pathway | 26 | 1 | 6.50E-03 |
| enzyme linked receptor protein signaling pathway | 35 | 1.4 | 6.70E-03 |
| neural tube formation | 8 | 0.3 | 6.70E-03 |
| monovalent inorganic cation homeostasis | 8 | 0.3 | 6.70E-03 |
| vasculature development | 27 | 1.1 | 6.90E-03 |
| heart development | 27 | 1.1 | 6.90E-03 |
| blood vessel development | 26 | 1 | 7.60E-03 |
| urogenital system development | 16 | 0.6 | 7.70E-03 |
| neural tube closure | 7 | 0.3 | 8.10E-03 |
| tube closure | 7 | 0.3 | 8.10E-03 |
| metal ion transport | 48 | 1.9 | 8.60E-03 |
| regulation of cell shape | 6 | 0.2 | 9.10E-03 |
| cellular ion homeostasis | 24 | 0.9 | 9.20E-03 |
| cellular chemical homeostasis | 24 | 0.9 | 9.20E-03 |
| glutamine family amino acid metabolic process | 10 | 0.4 | 1.20E-02 |
| embryonic epithelial tube formation | 8 | 0.3 | 1.50E-02 |
| tube lumen formation | 8 | 0.3 | 1.50E-02 |
| heart morphogenesis | 12 | 0.5 | 1.50E-02 |
| sodium ion transport | 15 | 0.6 | 1.50E-02 |
| organic acid catabolic process | 11 | 0.4 | 1.60E-02 |
| carboxylic acid catabolic process | 11 | 0.4 | 1.60E-02 |

| | | | |
|---|---|---|---|
| response to acid | 4 | 0.2 | 1.60E-02 |
| arginine biosynthetic process | 4 | 0.2 | 1.60E-02 |
| regulation of cell adhesion | 14 | 0.5 | 1.70E-02 |
| respiratory system development | 16 | 0.6 | 1.70E-02 |
| tRNA modification | 5 | 0.2 | 1.80E-02 |
| negative regulation of BMP signaling pathway | 5 | 0.2 | 1.80E-02 |
| cell-matrix adhesion | 10 | 0.4 | 2.10E-02 |
| ureteric bud development | 9 | 0.4 | 2.10E-02 |
| blood vessel morphogenesis | 21 | 0.8 | 2.20E-02 |
| nucleoside monophosphate metabolic process | 12 | 0.5 | 2.40E-02 |
| proteolysis | 81 | 3.2 | 2.40E-02 |
| cation transport | 57 | 2.2 | 2.50E-02 |
| mesenchymal cell differentiation | 10 | 0.4 | 2.70E-02 |
| mesenchyme development | 10 | 0.4 | 2.70E-02 |
| nitrogen compound biosynthetic process | 45 | 1.8 | 2.80E-02 |
| ribonucleoside monophosphate metabolic process | 7 | 0.3 | 2.80E-02 |
| ribonucleoside monophosphate biosynthetic process | 7 | 0.3 | 2.80E-02 |
| carboxylic acid transport | 12 | 0.5 | 2.90E-02 |
| organic acid transport | 12 | 0.5 | 2.90E-02 |
| response to carbohydrate stimulus | 5 | 0.2 | 3.10E-02 |
| response to glucose stimulus | 5 | 0.2 | 3.10E-02 |
| response to monosaccharide stimulus | 5 | 0.2 | 3.10E-02 |
| response to hexose stimulus | 5 | 0.2 | 3.10E-02 |
| amine transport | 11 | 0.4 | 3.10E-02 |
| homeostasis of number of cells | 13 | 0.5 | 3.20E-02 |
| branching morphogenesis of a tube | 13 | 0.5 | 3.20E-02 |
| alcohol catabolic process | 13 | 0.5 | 3.20E-02 |
| chloride transport | 10 | 0.4 | 3.40E-02 |
| regulation of cell morphogenesis | 10 | 0.4 | 3.40E-02 |
| cell-substrate adhesion | 10 | 0.4 | 3.40E-02 |
| forebrain development | 18 | 0.7 | 3.60E-02 |
| arginine metabolic process | 4 | 0.2 | 3.60E-02 |
| odontogenesis of dentine-containing tooth | 9 | 0.4 | 3.60E-02 |
| odontogenesis | 9 | 0.4 | 3.60E-02 |
| branching involved in ureteric bud morphogenesis | 6 | 0.2 | 3.70E-02 |
| negative regulation of cell cycle | 6 | 0.2 | 3.70E-02 |
| ureteric bud morphogenesis | 6 | 0.2 | 3.70E-02 |
| regulation of pH | 6 | 0.2 | 3.70E-02 |
| RNA modification | 8 | 0.3 | 3.80E-02 |
| kidney development | 13 | 0.5 | 3.80E-02 |
| gastrulation | 11 | 0.4 | 3.90E-02 |
| carbohydrate catabolic process | 14 | 0.5 | 4.10E-02 |
| glycolysis | 10 | 0.4 | 4.20E-02 |
| nucleoside monophosphate biosynthetic process | 10 | 0.4 | 4.20E-02 |
| morphogenesis of a branching structure | 13 | 0.5 | 4.50E-02 |
| integrin-mediated signaling pathway | 9 | 0.4 | 4.50E-02 |
| cell maturation | 9 | 0.4 | 4.50E-02 |
| metanephros development | 9 | 0.4 | 4.50E-02 |
| neural tube development | 11 | 0.4 | 4.70E-02 |
| regulation of BMP signaling pathway | 5 | 0.2 | 4.90E-02 |

| | | | |
|---|---|---|---|
| formation of primary germ layer | 8 | 0.3 | 4.90E-02 |
| somitogenesis | 8 | 0.3 | 4.90E-02 |
| inorganic anion transport | 12 | 0.5 | 5.00E-02 |

**Significantly detected KEGG pathways based on up-regulated DEGs between EGK.VIII and EGK.X (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Proteasome | 14 | 0.5 | 2.80E-03 |
| Caffeine metabolism | 5 | 0.2 | 4.30E-03 |
| Drug metabolism | 12 | 0.5 | 6.10E-03 |
| ECM-receptor interaction | 21 | 0.8 | 1.20E-02 |
| Starch and sucrose metabolism | 10 | 0.4 | 1.30E-02 |
| Glycerophospholipid metabolism | 17 | 0.7 | 1.60E-02 |
| Valine, leucine and isoleucine biosynthesis | 5 | 0.2 | 2.20E-02 |
| Purine metabolism | 30 | 1.2 | 3.40E-02 |

**Significantly detected BP terms of GO based on down-regulated DEGs between EGK.VIII and EGK.X (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| regulation of small GTPase mediated signal transduction | 41 | 1.6 | 4.80E-06 |
| regulation of Ras protein signal transduction | 35 | 1.3 | 6.60E-06 |
| phosphorus metabolic process | 116 | 4.4 | 3.20E-05 |
| phosphate metabolic process | 116 | 4.4 | 3.20E-05 |
| phosphorylation | 94 | 3.6 | 8.40E-04 |
| intracellular signaling cascade | 82 | 3.1 | 9.70E-04 |
| regulation of Ras GTPase activity | 17 | 0.6 | 9.80E-04 |
| regulation of Rab protein signal transduction | 13 | 0.5 | 1.00E-03 |
| regulation of Rab GTPase activity | 13 | 0.5 | 1.00E-03 |
| regulation of GTPase activity | 17 | 0.6 | 3.30E-03 |
| protein amino acid phosphorylation | 79 | 3 | 3.80E-03 |
| myeloid leukocyte activation | 7 | 0.3 | 5.40E-03 |
| vesicle-mediated transport | 34 | 1.3 | 8.30E-03 |
| dephosphorylation | 22 | 0.8 | 1.00E-02 |
| regulation of hydrolase activity | 24 | 0.9 | 1.20E-02 |
| reproductive structure development | 13 | 0.5 | 1.50E-02 |
| lipid localization | 14 | 0.5 | 1.60E-02 |
| membrane organization | 19 | 0.7 | 2.40E-02 |
| gonad development | 11 | 0.4 | 2.40E-02 |
| development of primary sexual characteristics | 13 | 0.5 | 2.70E-02 |
| two-component signal transduction system (phosphorelay) | 4 | 0.2 | 2.90E-02 |
| fatty acid metabolic process | 14 | 0.5 | 3.20E-02 |
| endocrine system development | 10 | 0.4 | 3.30E-02 |
| regulation of synaptic transmission | 8 | 0.3 | 3.40E-02 |
| negative regulation of programmed cell death | 25 | 1 | 3.60E-02 |
| negative regulation of cell death | 25 | 1 | 3.60E-02 |
| reproductive developmental process | 22 | 0.8 | 3.60E-02 |
| male gonad development | 7 | 0.3 | 3.70E-02 |
| regulation of programmed cell death | 44 | 1.7 | 3.70E-02 |
| regulation of Rho protein signal transduction | 13 | 0.5 | 3.70E-02 |
| regulation of cell death | 44 | 1.7 | 4.00E-02 |

| | | | |
|---|---|---|---|
| regulation of apoptosis | 43 | 1.6 | 4.20E-02 |
| regulation of transmission of nerve impulse | 8 | 0.3 | 4.30E-02 |
| regulation of neurological system process | 8 | 0.3 | 4.30E-02 |
| lipid transport | 12 | 0.5 | 4.30E-02 |
| multicellular organism reproduction | 22 | 0.8 | 4.40E-02 |
| reproductive process in a multicellular organism | 22 | 0.8 | 4.40E-02 |
| activation of protein kinase activity | 9 | 0.3 | 4.70E-02 |
| positive regulation of transferase activity | 16 | 0.6 | 4.80E-02 |
| negative regulation of apoptosis | 24 | 0.9 | 5.00E-02 |
| reproductive cellular process | 14 | 0.5 | 5.00E-02 |
| membrane invagination | 11 | 0.4 | 5.00E-02 |
| endocytosis | 11 | 0.4 | 5.00E-02 |
| microtubule cytoskeleton organization | 11 | 0.4 | 5.00E-02 |

**Significantly detected KEGG pathways based on down-regulated DEGs between EGK.VIII and EGK.X (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Glycerolipid metabolism | 14 | 0.5 | 7.00E-03 |
| Drug metabolism | 10 | 0.4 | 2.60E-02 |
| Oocyte meiosis | 23 | 0.9 | 3.80E-02 |
| Steroid biosynthesis | 6 | 0.2 | 4.20E-02 |

**Table 3-8. Gene list of functional annotations in heatmaps in Figure 3-3B**

**Cell Cycle**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000169 | PCNA | Proliferating cell nuclear antigen |
| ENSGALG00000000596 | CDC27 | cell division cycle protein 27 homolog |
| ENSGALG00000001277 | STAG1 | stromal antigen 1 |
| ENSGALG00000001332 | RBL1 | retinoblastoma-like 1 |
| ENSGALG00000001506 | PTTG1 | securin isoform 1 |
| ENSGALG00000001616 | CDC45 | cell division cycle 45 |
| ENSGALG00000001722 | E2F4 | transcription factor E2F4 |
| ENSGALG00000003045 | E2F1 | transcription factor E2F1 |
| ENSGALG00000003085 | CDK1 | cyclin-dependent kinase 1 |
| ENSGALG00000003297 | HDAC1 | Histone deacetylase 1 |
| ENSGALG00000003485 | CCND3 | G1/S-specific cyclin-D3 |
| ENSGALG00000003826 | ANAPC7 | anaphase promoting complex subunit 7 |
| ENSGALG00000004161 | CCNB2 | G2/mitotic-specific cyclin-B2 |
| ENSGALG00000004342 | ORC6 | origin recognition complex subunit 6 isoform 1 |
| ENSGALG00000004494 | CCNE1 | G1/S-specific cyclin-E1 |
| ENSGALG00000004640 | MAD2L2 | Mitotic spindle assembly checkpoint protein MAD2B |
| ENSGALG00000004838 | BUB1B | mitotic checkpoint serine/threonine-protein kinase BUB1 beta |
| ENSGALG00000004934 | CDC25A | M-phase inducer phosphatase 1 |
| ENSGALG00000006037 | MCM2 | DNA replication licensing factor MCM2 |
| ENSGALG00000006051 | CDC7 | cell division cycle 7 |
| ENSGALG00000006110 | PLK1 | serine/threonine-protein kinase PLK1 |
| ENSGALG00000007555 | CCND1 | G1/S-specific cyclin-D1 |
| ENSGALG00000007762 | CREBBP | CREB binding protein |

| | | |
|---|---|---|
| ENSGALG00000008233 | BUB1 | mitotic checkpoint serine/threonine-protein kinase BUB1 |
| ENSGALG00000008234 | ORC2 | origin recognition complex subunit 2 |
| ENSGALG00000008254 | ANAPC1 | anaphase promoting complex subunit 1 |
| ENSGALG00000008462 | CDK3 | cell division protein kinase 3 |
| ENSGALG00000008585 | SMC3 | structural maintenance of chromosomes protein 3 |
| ENSGALG00000008963 | DBF4 | DBF4 zinc finger |
| ENSGALG00000009612 | TGFB2 | transforming growth factor, beta 2 |
| ENSGALG00000009942 | MDM2 | Mdm2, transformed 3T3 cell double minute 2, p53 binding protein |
| ENSGALG00000009971 | CDC20 | cell division cycle protein 20 homolog |
| ENSGALG00000010623 | ORC1 | origin recognition complex subunit 1 |
| ENSGALG00000011881 | CCNA2 | cyclin A2 |
| ENSGALG00000011967 | MAD2L1 | mitotic spindle assembly checkpoint protein MAD2A |
| ENSGALG00000012348 | MCM6 | DNA replication licensing factor MCM6 |
| ENSGALG00000012402 | CUL1 | cullin 1 |
| ENSGALG00000012546 | MCM5 | DNA replication licensing factor MCM5 |
| ENSGALG00000012914 | PRKDC | DNA-dependent protein kinase catalytic subunit |
| ENSGALG00000012915 | MCM4 | minichromosome maintenance complex component 4 |
| ENSGALG00000014228 | SMC1B | structural maintenance of chromosomes 1B |
| ENSGALG00000014790 | CDK7 | cyclin-dependent kinase 7 |
| ENSGALG00000014991 | HDAC2 | histone deacetylase 2 |
| ENSGALG00000015641 | CCNH | cyclin H |
| ENSGALG00000015800 | ORC3 | origin recognition complex subunit 3 |
| ENSGALG00000015874 | TTK | TTK protein kinase |
| ENSGALG00000015985 | CCNE2 | G1/S-specific cyclin-E2 |
| ENSGALG00000016119 | RAD21 | double-strand-break repair protein rad21 homolog |
| ENSGALG00000016308 | MYC | Myc proto-oncogene protein |

| | | |
|---|---|---|
| ENSGALG00000016676 | MCM3 | DNA replication licensing factor MCM3 |
| ENSGALG00000016816 | CDC16 | cell division cycle protein 16 homolog |
| ENSGALG00000016823 | Unknown | Uncharacterized protein |
| ENSGALG00000017052 | CCNA1 | cyclin A1 |
| ENSGALG00000017283 | CCND2 | G1/S-specific cyclin-D2 |
| ENSGALG00000017358 | YWHAG | 14-3-3 protein gamma |

**DNA replication**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000169 | PCNA | Proliferating cell nuclear antigen |
| ENSGALG00000000734 | RPA2 | replication protein A 32 kDa subunit |
| ENSGALG00000004037 | DNA2 | DNA replication helicase/nuclease 2 |
| ENSGALG00000006037 | MCM2 | DNA replication licensing factor MCM2 |
| ENSGALG00000007665 | POLE | polymerase (DNA directed), epsilon, catalytic subunit |
| ENSGALG00000008703 | RFC4 | replication factor C subunit 4 |
| ENSGALG00000010700 | RPA3 | replication protein A3, 14kDa |
| ENSGALG00000012237 | POLE2 | DNA polymerase epsilon subunit 2 |
| ENSGALG00000012348 | MCM6 | DNA replication licensing factor MCM6 |
| ENSGALG00000012546 | MCM5 | DNA replication licensing factor MCM5 |
| ENSGALG00000012915 | MCM4 | minichromosome maintenance complex component 4 |
| ENSGALG00000014298 | RFC1 | replication factor C subunit 1 |
| ENSGALG00000016278 | PRIM2 | DNA primase large subunit |
| ENSGALG00000016312 | POLA1 | polymerase (DNA directed), alpha 1, catalytic subunit |
| ENSGALG00000016389 | RNASEH1 | ribonuclease H1 precursor |
| ENSGALG00000016676 | MCM3 | DNA replication licensing factor MCM3 |
| ENSGALG00000017064 | RFC3 | replication factor C subunit 3 |

| | | |
|---|---|---|
| ENSGALG00000017307 | POLD3 | DNA polymerase delta subunit 3 |
| ENSGALG00000024076 | FEN1 | Flap endonuclease 1 |

**MAPK signaling**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000525 | MAP3K3 | mitogen-activated protein kinase kinase kinase 3 |
| ENSGALG00000000534 | il-1beta | Gallus gallus interleukin 1, beta (IL1B), mRNA. |
| ENSGALG00000000600 | PTPN7 | protein tyrosine phosphatase, non-receptor type 7 |
| ENSGALG00000000681 | PAK1 | serine/threonine-protein kinase PAK 1 |
| ENSGALG00000001001 | MAP2K4 | mitogen-activated protein kinase kinase 4 |
| ENSGALG00000001267 | MAP2K2 | dual specificity mitogen-activated protein kinase kinase 2 |
| ENSGALG00000001475 | STMN1 | Stathmin |
| ENSGALG00000001501 | MAPK1 | mitogen-activated protein kinase 1 |
| ENSGALG00000002283 | MAPKAPK3 | mitogen-activated protein kinase-activated protein kinase 3 |
| ENSGALG00000003289 | CHUK | inhibitor of nuclear factor kappa-B kinase subunit alpha |
| ENSGALG00000003706 | DUSP1 | dual specificity protein phosphatase 1 |
| ENSGALG00000003845 | MKNK2 | MAP kinase interacting serine/threonine kinase 2 |
| ENSGALG00000004082 | TAOK1 | TAO kinase 1 |
| ENSGALG00000004098 | STK4 | serine/threonine-protein kinase 4 |
| ENSGALG00000004368 | GNA12 | guanine nucleotide binding protein (G protein) alpha 12 |
| ENSGALG00000004735 | MAP2K3 | dual specificity mitogen-activated protein kinase kinase 3 |
| ENSGALG00000004743 | MAPKAPK5 | MAP kinase-activated protein kinase 5 |
| ENSGALG00000004998 | RAF1 | RAF proto-oncogene serine/threonine-protein kinase |
| ENSGALG00000005400 | CACNA2D3 | calcium channel, voltage-dependent, alpha 2/delta subunit 3 |
| ENSGALG00000005660 | NF1 | neurofibromin 1 |
| ENSGALG00000005699 | NLK | nemo-like kinase |

| | | |
|---|---|---|
| ENSGALG00000005998 | RRAS2 | ras-related protein R-Ras2 |
| ENSGALG00000006109 | MAPK8 | mitogen-activated protein kinase 8 |
| ENSGALG00000006368 | PTPN5 | protein tyrosine phosphatase, non-receptor type 5 (striatum-enriched) |
| ENSGALG00000006390 | CRKL | v-crk avian sarcoma virus CT10 oncogene homolog-like |
| ENSGALG00000006426 | PAK2 | p21 protein (Cdc42/Rac)-activated kinase 2 |
| ENSGALG00000006508 | FGF13 | fibroblast growth factor 13 |
| ENSGALG00000006655 | MAP3K13 | mitogen-activated protein kinase kinase kinase 13 |
| ENSGALG00000006885 | H-RAS | GTPase HRas |
| ENSGALG00000007130 | cRac1A | ras-related C3 botulinum toxin substrate 1 |
| ENSGALG00000007356 | MAP3K8 | mitogen-activated protein kinase kinase kinase 8 |
| ENSGALG00000007623 | CACNA1G | calcium channel, voltage-dependent, T type, alpha 1G subunit |
| ENSGALG00000007687 | MAP2K1 | dual specificity mitogen-activated protein kinase kinase 1 |
| ENSGALG00000007706 | FGF8 | fibroblast growth factor 8 precursor |
| ENSGALG00000007806 | FGF16 | fibroblast growth factor 16 |
| ENSGALG00000007932 | TRAF6 | Uncharacterized protein |
| ENSGALG00000008016 | GRB2 | growth factor receptor-bound protein 2 |
| ENSGALG00000008058 | PAK3 | p21 protein (Cdc42/Rac)-activated kinase 3 |
| ENSGALG00000008569 | CHP1 | Calcineurin B homologous protein 1 |
| ENSGALG00000008581 | DUSP5 | dual specificity phosphatase 5 |
| ENSGALG00000008612 | MAPK11 | mitogen-activated protein kinase 11 |
| ENSGALG00000008970 | NRK | Nik related kinase |
| ENSGALG00000009287 | ATF2 | Cyclic AMP-dependent transcription factor ATF-2 |
| ENSGALG00000009344 | Unknown | Uncharacterized protein |
| ENSGALG00000009437 | MECOM | MDS1 and EVI1 complex locus |
| ENSGALG00000009450 | DUSP10 | dual specificity protein phosphatase 10 |
| ENSGALG00000009612 | TGFB2 | transforming growth factor, beta 2 |

| | | |
|---|---|---|
| ENSGALG00000010148 | PTPRR | protein tyrosine phosphatase, receptor type, R |
| ENSGALG00000010435 | RASGRP3 | ras guanyl-releasing protein 3 |
| ENSGALG00000010440 | MKNK1 | MAP kinase interacting serine/threonine kinase 1 |
| ENSGALG00000011419 | DUSP4 | Gallus gallus dual specificity phosphatase 4 (DUSP4), mRNA. |
| ENSGALG00000011620 | AKT1 | RAC-alpha serine/threonine-protein kinase |
| ENSGALG00000011917 | PPM1A | protein phosphatase, Mg2+/Mn2+ dependent, 1A |
| ENSGALG00000012055 | Unknown | Uncharacterized protein |
| ENSGALG00000012150 | TAB1 | TGF-beta-activated kinase 1 and MAP3K7-binding protein 1 |
| ENSGALG00000012163 | BDNF | Brain-derived neurotrophic factor |
| ENSGALG00000012273 | SOS2 | son of sevenless homolog 2 (Drosophila) |
| ENSGALG00000012280 | PPP3CA | protein phosphatase 3, catalytic subunit, alpha isozyme |
| ENSGALG00000012281 | PLA2G6 | 85 kDa calcium-independent phospholipase A2 |
| ENSGALG00000012304 | NFKB1 | Nuclear factor NF-kappa-B p105 subunit Nuclear factor NF-kappa-B p50 subunit |
| ENSGALG00000012356 | TAB2 | TGF-beta activated kinase 1/MAP3K7 binding protein 2 |
| ENSGALG00000013663 | FGF20 | fibroblast growth factor 20 |
| ENSGALG00000013761 | MAPK9 | Mitogen-activated protein kinase 9 |
| ENSGALG00000014631 | ACVR1B | activin A receptor, type IB |
| ENSGALG00000014872 | FGF10 | fibroblast growth factor 10 precursor |
| ENSGALG00000015406 | c-mos | Serine/threonine-protein kinase mos |
| ENSGALG00000015596 | MAP3K7 | mitogen-activated protein kinase kinase kinase 7 |
| ENSGALG00000016308 | MYC | Myc proto-oncogene protein |
| ENSGALG00000016406 | RPS6KA3 | ribosomal protein S6 kinase, 90kDa, polypeptide 3 |
| ENSGALG00000017706 | Unknown | Uncharacterized protein |
| ENSGALG00000019233 | Unknown | Uncharacterized protein |
| ENSGALG00000019384 | MAPK12 | mitogen-activated protein kinase 12 |
| ENSGALG00000019759 | MAPK14 | mitogen-activated protein kinase 14 |

| | | |
|---|---|---|
| ENSGALG00000020003 | MAP3K4 | mitogen-activated protein kinase kinase kinase 4 |
| ENSGALG00000020249 | RAPGEF2 | Rap guanine nucleotide exchange factor (GEF) 2 |
| ENSGALG00000021313 | PDGFRB | platelet-derived growth factor receptor, beta polypeptide |
| ENSGALG00000023648 | PPP3R1 | calcineurin subunit B type 1 |

**Table 3-9. Gene list of functional annotations in heatmaps in Figure 3-3C**

**Regulation of small GTPase mediated signal transduction**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000753 | TBC1D22B | TBC1 domain family, member 22B |
| ENSGALG00000001225 | RABGAP1 | RAB GTPase activating protein 1 |
| ENSGALG00000001244 | Unknown | Uncharacterized protein |
| ENSGALG00000001255 | ARHGEF37 | Rho guanine nucleotide exchange factor (GEF) 37 |
| ENSGALG00000001419 | DAB2IP | DAB2 interacting protein |
| ENSGALG00000001420 | RALGAPB | ral GTPase-activating protein subunit beta |
| ENSGALG00000001770 | ACAP3 | ArfGAP with coiled-coil, ankyrin repeat and PH domains 3 |
| ENSGALG00000001915 | VAV3 | guanine nucleotide exchange factor VAV3 |
| ENSGALG00000002699 | VAV2 | guanine nucleotide exchange factor VAV2 |
| ENSGALG00000002922 | NOTCH2 | neurogenic locus notch homolog protein 2 precursor |
| ENSGALG00000003052 | AGFG1 | ArfGAP with FG repeats 1 |
| ENSGALG00000003322 | TBC1D2B | TBC1 domain family, member 2B |
| ENSGALG00000003572 | Unknown | Uncharacterized protein |
| ENSGALG00000003712 | RAPGEF1 | Rap guanine nucleotide exchange factor (GEF) 1 |
| ENSGALG00000004323 | RASAL2 | RAS protein activator like 2 |
| ENSGALG00000004621 | PREX1 | phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1 |
| ENSGALG00000005157 | GIT2 | Gallus gallus G protein-coupled receptor kinase interactor 2 (GIT2), mRNA. |
| ENSGALG00000005781 | ARFGAP1 | ADP-ribosylation factor GTPase-activating protein 1 |
| ENSGALG00000005834 | PAFAH1B1 | Lissencephaly-1 homolog |
| ENSGALG00000005856 | TBC1D9B | TBC1 domain family, member 9B (with GRAM domain) |
| ENSGALG00000005868 | RAP1GAP2 | Rap1 GTPase-activating protein 2 |
| ENSGALG00000006130 | STMN3 | Stathmin-3 |
| ENSGALG00000006201 | TBC1D20 | TBC1 domain family, member 20 |

| | | |
|---|---|---|
| ENSGALG00000006439 | ARHGEF6 | Rho guanine nucleotide exchange factor 6 |
| ENSGALG00000006562 | MCF2 | MCF.2 cell line derived transforming sequence |
| ENSGALG00000006663 | ARHGEF12 | Rho guanine nucleotide exchange factor (GEF) 12 |
| ENSGALG00000006885 | H-RAS | GTPase HRas |
| ENSGALG00000006897 | Unknown | Uncharacterized protein |
| ENSGALG00000007040 | ACAP2 | Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2 |
| ENSGALG00000007093 | CYTH1 | cytohesin-1 |
| ENSGALG00000008247 | ARFGAP2 | ADP-ribosylation factor GTPase activating protein 2 |
| ENSGALG00000008600 | IQGAP1 | IQ motif containing GTPase activating protein 1 |
| ENSGALG00000008744 | Unknown | Unknown |
| ENSGALG00000009217 | TBC1D24 | Gallus gallus TBC1 domain family, member 24 (TBC1D24), mRNA. |
| ENSGALG00000009381 | SIPA1L1 | signal-induced proliferation-associated 1-like protein 1 |
| ENSGALG00000009855 | TBC1D9 | TBC1 domain family, member 9 (with GRAM domain) |
| ENSGALG00000010088 | RALGAPA1 | Ral GTPase activating protein, alpha subunit 1 (catalytic) |
| ENSGALG00000010435 | RASGRP3 | ras guanyl-releasing protein 3 |
| ENSGALG00000010903 | RASGEF1B | RasGEF domain family, member 1B |
| ENSGALG00000011251 | TBC1D5 | TBC1 domain family, member 5 |
| ENSGALG00000011342 | FGD6 | FYVE, RhoGEF and PH domain containing 6 |
| ENSGALG00000012273 | SOS2 | son of sevenless homolog 2 (Drosophila) |
| ENSGALG00000012975 | IQSEC3 | IQ motif and Sec7 domain 3 |
| ENSGALG00000012979 | TRIO | trio Rho guanine nucleotide exchange factor |
| ENSGALG00000013261 | ARHGEF11 | Rho guanine nucleotide exchange factor (GEF) 11 |
| ENSGALG00000013521 | TBC1D1 | TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1 |
| ENSGALG00000015273 | TBC1D23 | TBC1 domain family member 23 |
| ENSGALG00000015428 | PSD3 | pleckstrin and Sec7 domain containing 3 |
| ENSGALG00000015546 | TBC1D14 | TBC1 domain family member 14 |
| ENSGALG00000 | PREX2 | phosphatidylinositol-3,4,5-trisphosphate-dependent Rac |

| | | |
|---|---|---|
| 015564 | | exchange factor 2 |
| ENSGALG00000016006 | ITSN1 | intersectin 1 (SH3 domain protein) |
| ENSGALG00000016776 | TBC1D8 | TBC1 domain family, member 8 (with GRAM domain) |
| ENSGALG00000016817 | RASA3 | ras GTPase-activating protein 3 |
| ENSGALG00000017276 | ARHGEF5 | Rho guanine nucleotide exchange factor (GEF) 5 |
| ENSGALG00000017706 | Unknown | Uncharacterized protein |
| ENSGALG00000020249 | RAPGEF2 | Rap guanine nucleotide exchange factor (GEF) 2 |
| ENSGALG00000023703 | GIT1 | ARF GTPase-activating protein GIT1 |

**WNT signaling**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000889 | CSNK2A2 | casein kinase II subunit alpha' |
| ENSGALG00000001364 | CSNK1A1 | Casein kinase I isoform alpha |
| ENSGALG00000001600 | DVL1 | Uncharacterized protein |
| ENSGALG00000002588 | PPARD | peroxisome proliferator-activated receptor delta |
| ENSGALG00000002652 | FZD10 | frizzled-10 precursor |
| ENSGALG00000003485 | CCND3 | G1/S-specific cyclin-D3 |
| ENSGALG00000003767 | NKD1 | naked cuticle homolog 1 (Drosophila) |
| ENSGALG00000003916 | SIAH1 | siah E3 ubiquitin protein ligase 1 |
| ENSGALG00000004790 | WNT4 | Protein Wnt-4 |
| ENSGALG00000005410 | WNT5A | protein Wnt-5a |
| ENSGALG00000005699 | NLK | nemo-like kinase |
| ENSGALG00000006109 | MAPK8 | mitogen-activated protein kinase 8 |
| ENSGALG00000006480 | TCF1 | transcription factor 7 |
| ENSGALG00000006998 | LRP5 | low-density lipoprotein receptor-related protein 5 precursor |
| ENSGALG00000007130 | cRac1A | ras-related C3 botulinum toxin substrate 1 |
| ENSGALG00000007555 | CCND1 | G1/S-specific cyclin-D1 |
| ENSGALG00000 | Unknow | C-terminal binding protein-like |

| | | |
|---|---|---|
| 007642 | n | |
| ENSGALG00000007762 | CREBBP | CREB binding protein |
| ENSGALG00000007820 | BTRC | beta-transducin repeat containing E3 ubiquitin protein ligase |
| ENSGALG00000008414 | DVL3 | dishevelled segment polarity protein 3 |
| ENSGALG00000008433 | FZD2 | frizzled-2 precursor |
| ENSGALG00000008527 | PPP2R5D | protein phosphatase 2, regulatory subunit B', delta |
| ENSGALG00000008569 | CHP1 | Calcineurin B homologous protein 1 |
| ENSGALG00000008875 | PLCB1 | phospholipase C, beta 1 (phosphoinositide-specific) |
| ENSGALG00000009064 | FZD1 | frizzled-1 precursor |
| ENSGALG00000009556 | PRICKLE1 | prickle homolog 1 |
| ENSGALG00000010055 | DAAM2 | dishevelled associated activator of morphogenesis 2 |
| ENSGALG00000011260 | PPP2R5C | serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit gamma isoform |
| ENSGALG00000011355 | Wnt10a | protein Wnt-10a precursor |
| ENSGALG00000011358 | Wnt6 | Gallus gallus wingless-type MMTV integration site family, member 6 (WNT6), mRNA. |
| ENSGALG00000012280 | PPP3CA | protein phosphatase 3, catalytic subunit, alpha isozyme |
| ENSGALG00000012402 | CUL1 | cullin 1 |
| ENSGALG00000013366 | TBL1X | F-box-like/WD repeat-containing protein TBL1X |
| ENSGALG00000013761 | MAPK9 | Mitogen-activated protein kinase 9 |
| ENSGALG00000015025 | VANGL1 | VANGL planar cell polarity protein 1 |
| ENSGALG00000015596 | MAP3K7 | mitogen-activated protein kinase kinase kinase 7 |
| ENSGALG00000015718 | CTBP1 | C-terminal-binding protein 1 |
| ENSGALG00000016308 | MYC | Myc proto-oncogene protein |
| ENSGALG00000016451 | ROCK2 | Rho-associated, coiled-coil containing protein kinase 2 |
| ENSGALG00000016627 | FZD3 | frizzled-3 precursor |
| ENSGALG00000017283 | CCND2 | G1/S-specific cyclin-D2 |
| ENSGALG00000019315 | CSNK1E | casein kinase I isoform epsilon |
| ENSGALG00000023648 | PPP3R1 | calcineurin subunit B type 1 |

| | | |
|---|---|---|
| ENSGALG00000024120 | VANGL2 | VANGL planar cell polarity protein 2 |

**Notch signaling**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000001600 | DVL1 | Uncharacterized protein |
| ENSGALG00000001905 | DTX2 | deltex 2, E3 ubiquitin ligase |
| ENSGALG00000002055 | hairy1 | transcription factor HES-1 isoform 2 |
| ENSGALG00000002375 | NOTCH1 | notch 1 |
| ENSGALG00000002841 | RFNG | Beta-1,3-N-acetylglucosaminyltransferase radical fringe |
| ENSGALG00000002922 | NOTCH2 | neurogenic locus notch homolog protein 2 precursor |
| ENSGALG00000003297 | HDAC1 | Histone deacetylase 1 |
| ENSGALG00000003369 | KAT2A | K(lysine) acetyltransferase 2A |
| ENSGALG00000004284 | LFNG | beta-1,3-N-acetylglucosaminyltransferase lunatic fringe precursor |
| ENSGALG00000007642 | Unknown | C-terminal binding protein-like |
| ENSGALG00000007762 | CREBBP | CREB binding protein |
| ENSGALG00000008414 | DVL3 | dishevelled segment polarity protein 3 |
| ENSGALG00000008514 | DLL4 | delta-like 4 (Drosophila) |
| ENSGALG00000009020 | JAG1 | jagged 1 |
| ENSGALG00000009300 | NUMB | protein numb homolog |
| ENSGALG00000010501 | SNW1 | SNW domain containing 1 |
| ENSGALG00000010835 | DTX4 | deltex 4, E3 ubiquitin ligase |
| ENSGALG00000011696 | JAG2 | jagged 2 |
| ENSGALG00000012075 | DTX3L | deltex 3 like, E3 ubiquitin ligase |
| ENSGALG00000012442 | MFNG | MFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase |
| ENSGALG00000014276 | CIR1 | Corepressor interacting with RBPJ 1 |
| ENSGALG00000014365 | RBPJ | recombination signal binding protein for immunoglobulin kappa J region |
| ENSGALG00000014991 | HDAC2 | histone deacetylase 2 |

| | | |
|---|---|---|
| ENSGALG00000015718 | CTBP1 | C-terminal-binding protein 1 |

**Table 3-10. Gene list of functional annotations in heatmaps in Figure 3-5D**

**WNT signaling**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000839 | WNT11 | Protein Wnt-11 |
| ENSGALG00000001079 | Wnt3 | proto-oncogene Wnt-3 |
| ENSGALG00000001097 | WNT9b | Protein Wnt |
| ENSGALG00000002588 | PPARD | peroxisome proliferator-activated receptor delta |
| ENSGALG00000002652 | FZD10 | frizzled-10 precursor |
| ENSGALG00000003767 | NKD1 | naked cuticle homolog 1 (Drosophila) |
| ENSGALG00000003814 | Dkk-1 | Uncharacterized protein |
| ENSGALG00000004021 | AXIN2 | axin-2 |
| ENSGALG00000005123 | WNT7A | protein Wnt-7a |
| ENSGALG00000005396 | Wnt3a | Protein Wnt-3a |
| ENSGALG00000005401 | WNT9a | protein Wnt-9a precursor |
| ENSGALG00000005410 | WNT5A | protein Wnt-5a |
| ENSGALG00000005931 | RUVBL1 | ruvB-like 1 |
| ENSGALG00000006014 | PRKCB | protein kinase C, beta |
| ENSGALG00000006084 | Wnt8c | protein Wnt-8c precursor |
| ENSGALG00000006197 | CSNK2A1 | Casein kinase II subunit alpha |
| ENSGALG00000007332 | PRICKLE2 | prickle homolog 2 |
| ENSGALG00000007515 | SFRP5 | secreted frizzled-related protein 5 |
| ENSGALG00000007642 | Unknown | C-terminal binding protein-like |
| ENSGALG00000008433 | FZD2 | frizzled-2 precursor |
| ENSGALG00000008883 | TCF7L2 | transcription factor 7-like 2 |
| ENSGALG00000008907 | PLCB4 | 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta-4 |
| ENSGALG00000009320 | PSEN1 | Presenilin-1 Presenilin-1 NTF subunit Presenilin-1 CTF subunit |

| | | |
|---|---|---|
| ENSGALG00000009828 | PPP2R5A | protein phosphatase 2, regulatory subunit B', alpha |
| ENSGALG00000010529 | Lef-1 | lymphoid enhancer-binding factor 1 |
| ENSGALG00000011905 | CTNNB1 | catenin beta-1 |
| ENSGALG00000011992 | EP300 | E1A binding protein p300 |
| ENSGALG00000011993 | RBX1 | RING-box protein 1 |
| ENSGALG00000012036 | CAMK2D | calcium/calmodulin-dependent protein kinase type II delta chain |
| ENSGALG00000012456 | RAC2 | ras-related C3 botulinum toxin substrate 2 |
| ENSGALG00000012654 | NFATC1 | nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1 |
| ENSGALG00000012998 | WNT5B | protein Wnt-5b precursor |
| ENSGALG00000014235 | WNT7B | Protein Wnt-7b |
| ENSGALG00000014697 | MADH2 | mothers against decapentaplegic homolog 2 |
| ENSGALG00000014922 | ROCK1 | rho-associated protein kinase 1 |
| ENSGALG00000016069 | FZD6 | frizzled-6 precursor |
| ENSGALG00000016308 | MYC | Myc proto-oncogene protein |
| ENSGALG00000016627 | FZD3 | frizzled-3 precursor |
| ENSGALG00000016629 | PRKX | protein kinase, X-linked |
| ENSGALG00000023536 | FZD7 | frizzled-7 precursor |
| ENSGALG00000024120 | VANGL2 | VANGL planar cell polarity protein 2 |

**TGF-beta signaling**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000005198 | RPS6KB1 | ribosomal protein S6 kinase beta-1 |
| ENSGALG00000006158 | ACVR2B | activin receptor type-2B precursor |
| ENSGALG00000006210 | Id1 | DNA-binding protein inhibitor ID-1 |
| ENSGALG00000009256 | LEFTY2 | left-right determination factor 2 precursor |
| ENSGALG00000010346 | TGFB3 | transforming growth factor beta-3 preproprotein |
| ENSGALG00000011993 | RBX1 | RING-box protein 1 |

| | | |
|---|---|---|
| ENSGALG00000014697 | MADH2 | mothers against decapentaplegic homolog 2 |
| ENSGALG00000016308 | MYC | Myc proto-oncogene protein |
| ENSGALG00000016403 | ID2 | DNA-binding protein inhibitor ID-2 |

**Table 3-11. Gene list of functional annotations in heatmaps in Figure 3-5E**

**Neuron development**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000280 | EFNA5 | ephrin-A5 precursor |
| ENSGALG00000000644 | NFASC | neurofascin isoform 6 precursor |
| ENSGALG00000000936 | LMX1B | LIM/homeobox protein LMX-1.2 |
| ENSGALG00000002093 | NR2E3 | photoreceptor-specific nuclear receptor |
| ENSGALG00000002708 | LINGO1 | Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 1 |
| ENSGALG00000004631 | DRAXIN | Draxin |
| ENSGALG00000005256 | EPHA4 | Ephrin type-A receptor 4 |
| ENSGALG00000005396 | Wnt3a | Protein Wnt-3a |
| ENSGALG00000005730 | CELSR3 | cadherin, EGF LAG seven-pass G-type receptor 3 |
| ENSGALG00000006551 | tyrosine | tyrosine 3-monooxygenase |
| ENSGALG00000007140 | NRP1 | neuropilin-1 precursor |
| ENSGALG00000007668 | BMP7 | bone morphogenetic protein 7 |
| ENSGALG00000008411 | GHR | ghrelin/obestatin prepropeptide preproprotein |
| ENSGALG00000008494 | CREB1 | cyclic AMP-responsive element-binding protein 1 |
| ENSGALG00000008537 | EPHB3 | EPH receptor B3 |
| ENSGALG00000008930 | B3GNT2 | UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2 |
| ENSGALG00000009129 | DLX5 | Homeobox protein DLX-5 |
| ENSGALG00000009320 | PSEN1 | Presenilin-1 Presenilin-1 NTF subunit Presenilin-1 CTF subunit |
| ENSGALG00000009514 | NRCAM | neuronal cell adhesion molecule |
| ENSGALG00000011259 | RPE65 | Retinoid isomerohydrolase |
| ENSGALG00000011300 | TOP2B | DNA topoisomerase 2-beta |
| ENSGALG00000012357 | CXCR4 | C-X-C chemokine receptor type 4 |
| ENSGALG00000012462 | kif5c | Uncharacterized protein |

| | | |
|---|---|---|
| ENSGALG00000012594 | NTRK2 | neurotrophic tyrosine kinase, receptor, type 2 |
| ENSGALG00000012657 | SALL3 | sal-like protein 3 |
| ENSGALG00000013342 | GBX2 | Gallus gallus gastrulation brain homeobox 2 (GBX2), mRNA. |
| ENSGALG00000014419 | SLIT2 | slit homolog 2 protein precursor |
| ENSGALG00000015177 | GNA11 | guanine nucleotide-binding protein G(q) subunit alpha |
| ENSGALG00000015593 | EPHA7 | Ephrin type-A receptor 7 |
| ENSGALG00000015770 | APP | amyloid beta A4 protein precursor |
| ENSGALG00000020423 | NKX2-1 | NK2 homeobox 1 |

**Heart development**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000001416 | ADRA1B | adrenoceptor alpha 1B |
| ENSGALG00000001459 | TNNC1 | Troponin C, slow skeletal and cardiac muscles |
| ENSGALG00000001590 | PBRM1 | protein polybromo-1 |
| ENSGALG00000002546 | COL5A1 | collagen alpha-1(V) chain precursor |
| ENSGALG00000003209 | NODAL | nodal growth differentiation factor |
| ENSGALG00000004270 | ALDH1A2 | Retinal dehydrogenase 2 |
| ENSGALG00000005180 | COL11A1 | collagen, type XI, alpha 1 |
| ENSGALG00000005243 | PPP3CB | serine/threonine-protein phosphatase 2B catalytic subunit beta isoform |
| ENSGALG00000005396 | Wnt3a | Protein Wnt-3a |
| ENSGALG00000005448 | MYL3 | myosin light chain 1, cardiac muscle |
| ENSGALG00000006025 | XIRP1 | xin actin-binding repeat-containing protein 1 |
| ENSGALG00000006038 | TGFBR3 | transforming growth factor beta receptor type 3 precursor |
| ENSGALG00000006158 | ACVR2B | activin receptor type-2B precursor |
| ENSGALG00000006210 | Id1 | DNA-binding protein inhibitor ID-1 |
| ENSGALG00000006551 | tyrosine | tyrosine 3-monooxygenase |
| ENSGALG00000006594 | POFUT1 | GDP-fucose protein O-fucosyltransferase 1 precursor |

| | | |
|---|---|---|
| ENSGALG00000006629 | RBP4 | Retinol-binding protein 4 |
| ENSGALG00000006860 | KIF3A | kinesin-like protein KIF3A |
| ENSGALG00000007140 | NRP1 | neuropilin-1 precursor |
| ENSGALG00000007145 | ITGB1 | integrin beta-1 precursor |
| ENSGALG00000007706 | FGF8 | fibroblast growth factor 8 precursor |
| ENSGALG00000008148 | MYBPC3 | myosin binding protein C, cardiac |
| ENSGALG00000008621 | NRP2 | neuropilin 2 |
| ENSGALG00000008830 | BMP2 | bone morphogenetic protein 2 precursor |
| ENSGALG00000008978 | ITGA4 | integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor) |
| ENSGALG00000009320 | PSEN1 | Presenilin-1 Presenilin-1 NTF subunit Presenilin-1 CTF subunit |
| ENSGALG00000009844 | ACTC1 | Actin, alpha cardiac muscle 1 |
| ENSGALG00000010013 | EDNRA | endothelin-1 receptor precursor |
| ENSGALG00000010638 | CASP3 | caspase-3 |
| ENSGALG00000011442 | TGFBR2 | TGF-beta receptor type-2 |
| ENSGALG00000011630 | GLI2 | zinc finger protein GLI2 isoform 1 |
| ENSGALG00000011870 | HIF1A | hypoxia-inducible factor 1-alpha |
| ENSGALG00000011905 | CTNNB1 | catenin beta-1 |
| ENSGALG00000011951 | RPSA | 40S ribosomal protein SA |
| ENSGALG00000012329 | GLI3 | transcriptional activator GLI3 |
| ENSGALG00000012429 | BMP4 | bone morphogenetic protein 4 precursor |
| ENSGALG00000012617 | ChALK5 | TGF-beta receptor type-1 precursor |
| ENSGALG00000012620 | PTCH1 | protein patched homolog 1 |
| ENSGALG00000013781 | MYH7 | Gallus gallus myosin, heavy chain 7, cardiac muscle, beta (MYH7), mRNA. |
| ENSGALG00000013818 | CITED2 | cbp/p300-interacting transactivator 2 |
| ENSGALG00000014976 | GATA6 | GATA binding protein 6 |
| ENSGALG00000015013 | MSX1 | Homeobox protein GHOX-7 |
| ENSGALG000 | GNA11 | guanine nucleotide-binding protein G(q) subunit alpha |

| | | |
|---|---|---|
| 00015177 | | |
| ENSGALG00000015935 | SMYD1 | SET and MYND domain-containing protein 1 |
| ENSGALG00000017372 | SOX14 | transcription factor SOX-14 |

**Lung development**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000002203 | FGF18 | Gallus gallus fibroblast growth factor 18 (FGF18), mRNA. |
| ENSGALG00000002909 | CTGF | connective tissue growth factor precursor |
| ENSGALG00000003209 | NODAL | nodal growth differentiation factor |
| ENSGALG00000004270 | ALDH1A2 | Retinal dehydrogenase 2 |
| ENSGALG00000005410 | WNT5A | protein Wnt-5a |
| ENSGALG00000006158 | ACVR2B | activin receptor type-2B precursor |
| ENSGALG00000006378 | LIPA | lipase A, lysosomal acid, cholesterol esterase |
| ENSGALG00000006629 | RBP4 | Retinol-binding protein 4 |
| ENSGALG00000007706 | FGF8 | fibroblast growth factor 8 precursor |
| ENSGALG00000009327 | SP3 | Transcription factor Sp3 |
| ENSGALG00000009495 | FGFR2 | Fibroblast growth factor receptor 2 |
| ENSGALG00000010005 | EPAS1 | endothelial PAS domain-containing protein 1 |
| ENSGALG00000010290 | VEGFA | Vascular endothelial growth factor A |
| ENSGALG00000011442 | TGFBR2 | TGF-beta receptor type-2 |
| ENSGALG00000011905 | CTNNB1 | catenin beta-1 |
| ENSGALG00000012329 | GLI3 | transcriptional activator GLI3 |
| ENSGALG00000012429 | BMP4 | bone morphogenetic protein 4 precursor |
| ENSGALG00000013907 | KDR | vascular endothelial growth factor receptor 2 precursor |
| ENSGALG00000013929 | PDGFRA | platelet-derived growth factor receptor alpha precursor |
| ENSGALG00000014976 | GATA6 | GATA binding protein 6 |
| ENSGALG00000016906 | SPRY2 | protein sprouty homolog 2 |

| | | |
|---|---|---|
| ENSGALG00000017295 | PTHrP | Parathyroid hormone-related protein Osteostatin |
| ENSGALG00000018733 | Unknown | Unknown |
| ENSGALG00000020423 | NKX2-1 | NK2 homeobox 1 |

**ECM-receptor interaction**

| Ensembl_ID | Gene symbol | Description |
|---|---|---|
| ENSGALG00000000379 | ITGB3 | integrin beta-3 precursor |
| ENSGALG00000000569 | SDC3 | syndecan-3 precursor |
| ENSGALG00000001343 | LAMB3 | laminin, beta 3 |
| ENSGALG00000002041 | AGRN | agrin |
| ENSGALG00000002546 | COL5A1 | collagen alpha-1(V) chain precursor |
| ENSGALG00000002552 | COL3A1 | Gallus gallus collagen, type III, alpha 1 (COL3A1), mRNA. |
| ENSGALG00000002655 | ITGAV | integrin alpha-V precursor |
| ENSGALG00000003283 | COMP | cartilage oligomeric matrix protein |
| ENSGALG00000003578 | fn1 | fibronectin precursor |
| ENSGALG00000003589 | VTN | vitronectin precursor |
| ENSGALG00000003923 | COL6A3 | collagen alpha-3(VI) chain precursor |
| ENSGALG00000004627 | LAMC2 | laminin, gamma 2 |
| ENSGALG00000005022 | GP9 | glycoprotein IX (platelet) |
| ENSGALG00000005180 | COL11A1 | collagen, type XI, alpha 1 |
| ENSGALG00000006126 | COL6A2 | Collagen alpha-2(VI) chain |
| ENSGALG00000006859 | SV2B | synaptic vesicle glycoprotein 2B |
| ENSGALG00000007145 | ITGB1 | integrin beta-1 precursor |
| ENSGALG00000007849 | CD44 | CD44 antigen precursor |
| ENSGALG00000007905 | LAMB1 | laminin, beta 1 |
| ENSGALG00000007917 | LAMB4 | laminin, beta 4 |
| ENSGALG00000008978 | ITGA4 | integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor) |

| | | |
|---|---|---|
| ENSGALG00000009626 | THBS1 | thrombospondin-1 precursor |
| ENSGALG00000009641 | COL1A2 | collagen alpha-2(I) chain precursor |
| ENSGALG00000011778 | ITGB5 | integrin beta-5 precursor |
| ENSGALG00000014615 | laminin | laminin subunit alpha-1 precursor |
| ENSGALG00000015355 | CD47 | leukocyte surface antigen CD47 precursor |
| ENSGALG00000016480 | SDC1 | syndecan 1 |
| ENSGALG00000016841 | COL4A1 | collagen alpha-1(IV) chain precursor |
| ENSGALG00000016843 | COL4A2 | collagen, type IV, alpha 2 precursor |
| ENSGALG00000017272 | VWF | von Willebrand factor |
| ENSGALG00000019761 | CHAD | chondroadherin |

**Table 3-12. mRNAs used in the co-expression network analysis**

| Ensembl_ID | Gene symbol | Type | Function |
|---|---|---|---|
| ENSGALG00000009020 | JAG1 | Notch | Ligand |
| ENSGALG00000011696 | JAG2 | Notch | Ligand |
| ENSGALG00000008514 | DLL4 | Notch | Ligand |
| ENSGALG00000002375 | NOTCH1 | Notch | Receptor |
| ENSGALG00000002922 | NOTCH2 | Notch | Receptor |
| ENSGALG00000014365 | RBPJ | Notch | DNA binding |
| ENSGALG00000006885 | H-RAS | MAPK | Kinase |
| ENSGALG00000004998 | RAF1 | MAPK | Kinase |
| ENSGALG00000007687 | MAP2K1 | MAPK | Kinase |
| ENSGALG00000001267 | MAP2K2 | MAPK | Kinase |
| ENSGALG00000001501 | MAPK1 | MAPK | Kinase |
| ENSGALG00000001001 | MAP2K4 | MAPK | Kinase |
| ENSGALG00000006109 | MAPK8 | MAPK | Kinase |
| ENSGALG00000004735 | MAP2K3 | MAPK | Kinase |
| ENSGALG00000019759 | MAPK14 | MAPK | Kinase |
| ENSGALG00000004790 | WNT4 | Wnt | Ligand |
| ENSGALG00000006488 | WNT6 | Wnt | Ligand |
| ENSGALG00000008485 | FZD1 | Wnt | Receptor |
| ENSGALG00000011358 | RYK | Wnt | Receptor |
| ENSGALG00000009064 | CCNYL1 | Wnt | Receptor |
| ENSGALG00000011905 | WNT5B | Wnt | Ligand |
| ENSGALG00000003814 | WNT8C | Wnt | Ligand |
| ENSGALG00000006084 | DKK1 | Wnt | Ligand |
| ENSGALG00000012998 | FZD7 | Wnt | Receptor |
| ENSGALG00000023536 | FZD10 | Wnt | Receptor |
| ENSGALG00000002652 | CTNNB1 | Wnt | DNA binding |
| ENSGALG00000010529 | Lef-1 | Wnt | DNA binding |
| ENSGALG00000008883 | TCF7L2 | Wnt | DNA binding |
| ENSGALG00000010346 | TGFB3 | TGF-beta | Ligand |
| ENSGALG00000009256 | LEFTY2 | TGF-beta | Ligand |
| ENSGALG00000006158 | ACVR2B | TGF-beta | Receptor |
| ENSGALG00000014697 | MADH2 | TGF-beta | DNA binding |
| ENSGALG00000016308 | MYC | TGF-beta | DNA binding |
| ENSGALG00000006210 | Id1 | TGF-beta | DNA binding |
| ENSGALG00000016403 | ID2 | TGF-beta | DNA binding |
| ENSGALG00000016059 | ETS2 | TFs | DNA binding |
| ENSGALG00000008848 | SOX2 | TFs | DNA binding |
| ENSGALG00000027772 | NANOG | TFs | DNA binding |
| ENSGALG00000016308 | MYC | TFs | DNA binding |
| ENSGALG00000001143 | ETS1 | TFs | DNA binding |
| ENSGALG00000013956 | c-myb | TFs | DNA binding |
| ENSGALG00000026574 | POUV | TFs | DNA binding |

**Table 3-13. mRNAs used in orthologous gene expression comparison**

| Ensembl_Chicken | Ensembl_Human | Ensembl_Mouse | Gene symbol | Type |
|---|---|---|---|---|
| ENSGALG00000009020 | ENSG00000101384 | ENSMUSG00000027276 | JAG1 | Notch |
| ENSGALG00000011696 | ENSG00000184916 | ENSMUSG00000002799 | JAG2 | Notch |
| ENSGALG00000008514 | ENSG00000128917 | ENSMUSG00000027314 | DLL4 | Notch |
| ENSGALG00000002375 | ENSG00000148400 | ENSMUSG00000026923 | NOTCH1 | Notch |
| ENSGALG00000002922 | ENSG00000134250 | ENSMUSG00000027878 | NOTCH2 | Notch |
| ENSGALG00000014365 | ENSG00000168214 | ENSMUSG00000039191 | RBPJ | Notch |
| ENSGALG00000001501 | ENSG00000100030 | ENSMUSG00000063358 | MAPK1 | MAPK |
| ENSGALG00000001001 | ENSG00000065559 | ENSMUSG00000033352 | MAP2K4 | MAPK |
| ENSGALG00000006109 | ENSG00000107643 | ENSMUSG00000021936 | MAPK8 | MAPK |
| ENSGALG00000004735 | ENSG00000034152 | ENSMUSG00000018932 | MAP2K3 | MAPK |
| ENSGALG00000019759 | ENSG00000112062 | ENSMUSG00000053436 | MAPK14 | MAPK |
| ENSGALG00000006488 | ENSG00000115596 | ENSMUSG00000033227 | WNT6 | Wnt |
| ENSGALG00000011358 | ENSG00000163785 | ENSMUSG00000032547 | RYK | Wnt |
| ENSGALG00000011905 | ENSG00000111186 | ENSMUSG00000030170 | WNT5B | Wnt |
| ENSGALG00000002652 | ENSG00000168036 | ENSMUSG00000006932 | CTNNB1 | Wnt |
| ENSGALG00000010529 | ENSG00000138795 | ENSMUSG00000027985 | Lef-1 | Wnt |
| ENSGALG00000010346 | ENSG00000119699 | ENSMUSG00000021253 | TGFB3 | TGFbeta |
| ENSGALG00000006210 | ENSG00000125968 | ENSMUSG00000042745 | Id1 | TGFbeta |

**Table 3-14. Primer sequences used in the in situ hybridization**

| Gene | Primer sequences | |
|---|---|---|
| | Forward (5'→3') | Reverse (5'→3') |
| *WNT4* | AAGGGGCATCTTCCAACAGA | GGTTTGCACTTGACAGAGCA |
| *WNT6* | CAGCCTTCGTGTATGCCATC | CCTTGAAGGCTCCGTTGAAG |
| *NOTCH1* | CTGATCCTGGCTGCTCGCTT | CCTCCGGGCTTTCAGGTCTT |
| *NOTCH2* | CGTGTGAACACGCTGGGAAA | AGCCATTGGGATGGTCGATG |

## 4. Discussion

From the moment of fertilization, precise regulation of gene expression and dynamic courses occurs during early embryogenesis in various organisms. However, most of the early developmental programs in avian species have not been determined. In this study, RNA sequencing helped unravel the steps occurring in pre-oviposited embryonic development in the chicken. To examine systematically genetic regulation of chicken during early embryogenesis, in comparison with other species, transcriptomic data were analyzed from various perspectives.

Of many observations from our transcriptomic analysis, the most unique point is a very large number of actively expressed transcripts between the oocyte and zygote stages. This phenomenon occurs between the oocyte and zygote stages in other mammals, the so-called minor ZGA. In previous research, transcriptional activity was measured in terms of bromouridine triphosphate incorporation, and greater incorporation was seen in the male pronucleus from the 1-cell embryo stage (Aoki et al., 1997). Also, soon after fertilization in the mouse, epigenetic modification of paternal pronucleus alters vigorously (Aoshima et al., 2015). From this evidence, we specified this event the 1$^{st}$ ZGA, which remains to be elucidated deeply. Surprisingly, no transcriptional differences were observed among the zygote, EGK.I and EGK.III stages. Indeed, these results suggested the transcriptional silencing during early chicken embryonic development, as cleavage stages of human, from the 1- to 4-cell stages (Braude et al., 1988, Xue et al., 2013). Even though central cells positive for phosphorylation of the RNA polymerase II C-terminal domain (p-PolII) were apparent during the late EGK.II to early EGK.III stage according to a previous report (Nagai et al., 2015), significant expression were not found at this point when cells still divide rapidly, which

were confirmed at the transcriptome level in this study. In contrast, from EGK.III to EGK.VI stage, 1,287 transcripts with biological functions involved in transcription were upregulated, and only 13 transcripts were downregulated; this difference indicated the presence of another ZGA. It may be accompanied by lengthening of the cell cycle, like as *Xenopus* (Newport and Kirschner, 1982). Based on these observations, we suggest that 1st and 2nd ZGA events in the actual transcript level occurred between oocyte and zygote stages and between EGK.III and EGK.VI stages, respectively.

The 1st and 2nd ZGA induced genes seemed to be correlated functionally with the cleavage and area pellucida formation periods, respectively. From the functional network analysis results, we found the relevance between the early developmental courses and transcriptomic signatures. First, the genes activated during the 1st ZGA may be largely associated with the cleavage period. During these stages, we found several functional processes enriched after fertilization. Cell cycle, DNA replication and MAPK signaling appear to be involved in mitosis in early embryos. In sea urchin embryos after fertilization, the MAPK pathway, via cdk2 activation, regulates DNA replication and entry into mitosis (Zhang et al., 2005, Kisielewska et al., 2009). The cleavage stages during intrauterine development in the chicken appeared with rapid cellularization from a single cell to approximately 8,000 cells, within 10 h after the first cleavage event. Mitotic modulators from the 1st ZGA may regulate sustainable and rapid cell division in the early chicken embryo.

Notch signaling plays important roles in cell−cell interactions, mediated by Delta ligands and Notch receptors. As an example, in *C. elegans*, Notch signaling is involved in early cell fate specification for the asymmetric

differentiated cell functions (Priess, 2005). Similarly, chicken *NOTCH1* and *NOTCH2* were expressed asymmetrically in cleavage stage. Further investigation of the Notch signaling pathway during the 1st ZGA is needed to clarify whether it actually determines cell fate in the early embryo of chickens. Moreover, the spatiotemporal expression of marker genes, which are related with three germ layer in later stages, could be investigated to find the early fate mapping during cleavage stages in chicken. Wnt signaling and small GTPase signaling may interact to regulate cell polarity and cleavage patterns. During embryogenesis, maternal Wnt signaling is necessary for axis formation via the canonical pathway in *Xenopus* (Tao et al., 2005). In addition, a recent finding suggests that the non-canonical Wnt/STOP pathway of maternal Wnt signaling plays a role in the progression of cell cleavage during early *Xenopus* embryogenesis (Huang et al., 2015). Additionally, small GTPase RhoA acts as a regulator of spindle formation and cytokinesis during early embryo development in *C. elegans* and *Sus scrofa* (Tse et al., 2012, Zhang et al., 2014b). Furthermore, the Wnt/PCP pathway, via RhoA interaction with the actin cytoskeleton, regulates the orientation of cell division (Castanon et al., 2013). During early cleavage events in chicken, formation of the cleavage furrow in the embryonic central region was asymmetrical, in a radial manner (Lee et al., 2013). Moreover, layer increase through the cleavage period seems to be caused by cell division orientation, to some extent (Nagai et al., 2015). In this regard, maternal Wnt ligands in chicken, such as *WNT4* and *WNT6*, which differ from maternal *WNT11* in *Xenopus* (Tao et al., 2005), may regulate cleavage formation and orientation during early development. During cleavage in EGK.III, *WNT4* and *WNT6* showed asymmetric expression, suggesting the diversification of embryonic cells in this period. On the other hand, *WNT4* and *WNT6* would be further elucidated in terms of interaction with the receptor, due to their characteristic

on the species-specific maternal Wnt ligands. Consequently, the representative processes during cleavage period such as rapid asymmetric cellularization and layer increase could be governed by 1st ZGA genes.

Otherwise, those maternal Wnt ligands were downregulated and replaced with *WNT5B*, *WNT8C* and antagonist *DKK-1*, which could regulate axis formation, induced by the 2nd ZGA. In addition, TGF-beta signaling could also cooperate with Wnt signaling for the axis formation and lineage specification. Even though the interaction of Wnt and TGF-beta/Activin signaling is known to regulate initiation of primitive streak formation during post-ovipositional development (Skromne and Stern, 2001), their synergism in pre-ovipositional development has to be elucidated, mostly with higher expressed Wnt ligands including *WNT11*, *WNT3A*, *WNT5A*, *WNT9A*, in terms of both canonical and non-canonical pathway. Along with these signaling, we found the simultaneous expression of three germ layers such as neuronal development (ectoderm), heart development (mesoderm) and lung development (endoderm), after the EGK.VI stage by the functional enrichment analysis as well. Moreover, the ECM-receptor interaction pathway, which could plays an important role in tissue and organ morphogenesis, was identified during the area pellucida formation stage. Four specification markers, *SOX3* as an epiblast marker, and *SOX17*, *GATA4* and *GATA6* as endoderm markers, were identified in stages EGK.VII to VIII of the zebra finch (Mak et al., 2015), indicating molecular specification prior to morphological segregation. However, the simultaneous expression associated with three germ layer is hard to demonstrate the lineage specification during intrauterine development in chicken, which need to be studied further. Collectively, anterior-posterior axis formation and lineage segregation into hypoblast and epiblast seems to be prepared in area pellucida formation

period from EGK.VI stage by 2nd ZGA.

Finally, we compared gene expression patterns in chickens with other amniotes, humans and mice, during early embryogenesis. The ontogenetic physiology of early development in aves and mammals, such as cell division patterns and lineage segregation, seems to be quite different (Sheng, 2014). Even though some signaling genes were shown to be restricted to chicken development, unexpectedly, all stages of the chicken showed high similarity to the mammalian stages gradually after the minor and major ZGA, which occur in the pronucleus and 8-cell stage in humans and the 2-cell stage in mice, respectively (Lee et al., 2014). Furthermore, similar pattern were observed between chicken and zebrafish from the perspective of ZGA. Additionally, the relative similarities greatly increased after major ZGA than after minor ZGA, indicating the earliest genes of 1st ZGA or minor ZGA were evolutionarily different as in the comparison of other species (Heyn et al., 2014). Also, this observation suggests that the global transcriptome and basic cellular functions including cell cycle, small GTPase signaling, transcription, and translation in avian and mammalian embryos increasingly overlap and resemble each other over the course of development, even though the divergence in activation timing of them. In human and mouse, the fundamental programs such as cell cycle before major ZGA (4-cell in human and 1-cell in mouse), transcription after major ZGA (8-cell in human and 2-cell in mouse), and translation (morula in human and 8-cell and morula in mouse) were found in a previous study (Xue et al., 2013). In this study, cell cycle before EGK.VI, transcription after EGK.VI, and translation after EGK.VIII were enriched based on functional annotation. Considering all the available evidence, we summarized the chicken pre-oviposited developmental stages and correlated the chicken stages with those of early mammalian early

development, although there are differential gene expression related to signaling pathways between chicken and mammals. In conclusion, we report the first temporal transcriptomic analysis during early chicken development. Our results give an explanation of the developmental signaling features governed by two waves of ZGAs during early developmental events in avian species, at transcriptome resolution. From these results, we expect important studies in early avian development, including the fundamental transcriptome and systematic signaling pathways, from an evo-devo perspective. Furthermore, our results will contribute to understanding of early avian embryogenesis and to filling the gap for comparative studies with other amniotes in various developmental respects.

# CHAPTER 4

# Transcriptional and Translational Dynamics during Maternal-to-Zygotic Transition in Early Chicken Development

## 1. Introduction

Embryogenesis is well-coordinated developments with the predetermined genetic programs. Fundamental cellular processes such as transcription and translation govern the initial and critical events of early embryonic development after fertilization. During the first round of early development, a rapid and reductive cellularization occurs, which is mostly handled by the maternal mRNAs and proteins stored in the oocyte (Tadros and Lipshitz, 2009). Through the early mitotic divisions, the transcriptional inactivation is sustained until zygotic genome activation (ZGA) accompanied by lengthening of the cell cycle (Newport and Kirschner, 1982).

The massive induction of zygotic transcription and degradation of maternal mRNAs happen in the period called maternal-to-zygotic transition (MZT), which is a crucial process for establishing an individual organism (Lee et al., 2014). Another term, midblastula transition (MBT) which is originally defined in amphibian development, is a morphological embodiment of MZT and applies to non-mammalian species with respect to ZGA (Lee et al., 2014). The timing of ZGA and MZT varies among vertebrate species (Tadros and Lipshitz, 2009). The major ZGA in human and mouse occur in 2-cell and 8-cell stage, respectively (Xue et al., 2013, Yan et al., 2013). In zebrafish and frog, 128-cell stage has been reported as ZGA. It is well known that ZGA is triggered by maternal transcription factors such as Nanog, Pou5f1 and SoxB1 family deposited in oocyte (Lee et al., 2013, Leichsenring et al., 2013). Along with zygotic transcription, the conserved zygotic microRNAs (miRNAs) such as miR-430 in zebrafish, miR-427 in frog, and miR-290 in mouse are induced by these factors and have the role in maternal RNA degradation via post-transcriptional regulation in vertebrates (Giraldez et al.,

2006, Tang et al., 2007, Lund et al., 2009, Svoboda and Flemr, 2010, Yartseva and Giraldez, 2015). In addition, the variable expression of ribosomal RNA (rRNA) suggested that the translational dynamics such as limited translation and transition of maternal and zygotic ribosome pools in MZT (Gilbert et al., 2009, Ihara et al., 2011).

The intrauterine development of avian species comprises the two definite periods including cleavage and area pellucida formation, according to Eyal-Giladi and Kochav (EGK) criteria (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980). The first developmental transition equivalent to MBT indicating slower rate of cell cycle and morphogenesis occurs during area pellucida formation (Park et al., 2006). However, the molecular basis in the induction of zygotic developmental program, clearance of maternal factors, and post-transcriptional regulators during MZT has not been studied in avian species to date.

In this study, we provided the developmental transition in chicken according to transcriptional and translational-related aspects. Based on RNA-seq data of early chicken embryos, from oocyte to EGK.X, co-expression analysis was performed to analyze transcriptomic changes during the early developmental process. In addition, we tried to identify epigenetic dynamics including the conserved miRNA related to the transitional point in chicken. Finally, combining the identified transition and various biological knowledge, we attempted to establish the connection between well-described vertebrates and avian species in evolutionarily intermediate position.

## 2. Materials and methods

### *Experimental animals and animal care*

The care and experimental use of chickens were approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1). Chickens were maintained according to a standard management program at the University Animal Farm, Seoul National University, Korea. The procedures for animal management, reproduction and embryo manipulation adhered to the standard operating protocols of our laboratory.

### *RNA-seq data availability and pipeline for getting gene expression levels*

Our previously generated RNA-seq data were secondary employed, which was deposited in the GEO database (GSE86592). The experimental use of chickens was approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1). A total of 21 raw-sequencing data (biologically triplicated for each embryonic stage) were pre-processed using Trimmomatic (ver. 0.32) (Bolger et al., 2014). The number of embryos used for the biological triplicates of raw-sequencing data were: 5 oocytes for each replication; 3 zygotes for each replication; 6, 5, and 6 EGK.I embryos for each replication; 6 EGK.III embryos for each replication; 6, 6, and 7 EGK.VI embryos for each replication; 5, 5, and 3 EGK.VIII embryos for each replication; and 10 EGK.X embryos for each replication. After that, such clean reads were mapped on the chicken's reference genome (Ensembl; galGal4) using HISAT2 (Kim et al., 2015). Python script, HTSeq-count (Anders et al., 2014), was employed to quantify mapped reads in each Ensembl transcript and multidimensional scaling (MDS) method was used for the relationship among the samples (Figure 4-1A).

***Clustering analysis to identify co-expressed genes across all developmental stages***

A weighted correlation network analysis (WCGNA) (Langfelder and Horvath, 2008) was used to categorize gene expression patterns. As a result, eight representative modules were detected in the analysis (Figure 4-1B). Of them, seven representative patterns (Cluster 1~7) were finally observed after repetitive patterns were removed (Figure 4-1C); removal was based on the patterns observed in the WCGNA and the patterns expected from our biological assumption. In addition to seven representative patterns, the negatively correlated patterns (Cluster -1~-7) with Cluster 1~7 were included to develop additional clusters. To detect genes corresponding to the seven representative patterns, supervised learning was performed recursively to detect co-expressed genes conservatively. First, to classify genes into the seven patterns (Figure 4-1C), class labels were defined as follows:

Cluster 1 = (1, 0, 0, 0, 0, 1, 1)

Cluster 2 = (0, 1, 1, 1, 1, 1, 1)

Cluster 3 = (1, 2, 2, 2, 2, 1, 0)

Cluster 4 = (0, 0, 0, 0, 1, 1, 1)

Cluster 5 = (0, 0, 0, 0, 1, 2, 2)

Cluster 6 = (0, 0, 0, 0, 0, 1, 1)

Cluster 7 = (0, 1, 1, 1, 2, 2, 2)

The seven "0"-based vectors were defined to express the degree of fluctuation across the developmental stages, from oocyte to EGK.X. Based on these class labels, a decision tree-based classification analysis was performed. $Log_2$ trimmed mean of M-value (TMM) normalized values were used for a gene expression matrix to consider library size in each sample, calculated using edgeR (Robinson et al., 2010). Spearman's correlation coefficients were used for a distance matrix to characterize the linear relationship between class-label and gene-expression patterns.

$$r_{gc} = 1 - \frac{6\sum(ran(TMM_g), ran(Cluster_c))}{n(n^2-1)} \qquad \text{(Eq. 1)}$$

,where *ran(X)* represents the rank of observed values, *g* represents TMM normalized values for a specific gene, *c* represents the class label, and *n* represents the number of observations (7 (stages) × 3 (biological replications) = 21). To test their significance, Spearman's correction test was used, as follows:

$$\mathrm{t} = r_{gc}\sqrt{\frac{n-2}{1-r_{gc}^2}} \sim t(n\text{-}2) \qquad \text{(Eq. 2)}$$

Under the null hypothesis,

$H_0$: no linear relationship between c and g , (Eq. 2)

the statistic approximately follows a Student's *t* distribution. Based on this statistic, a permutation test was performed (Kendall et al., 1968), with

a false discovery rate (FDR) adjusted P < 0.1 considered as the threshold for a significant relationship. From these class labels ($Class_c$), the correlation-based distance matrix ($C_{gc}$) and decision matrix of significance test ($S_{gc}$, 0 and 1 represent significant and not significant, respectively), decision tree-based classification was performed as follows (**Pseudo code 1**):

**[Pseudo code 1]**

Classification matrix denotes $D_g$ for each gene $g$

**For** *i* **in** 1 **to** $g$

**If**$(\sum_{c=1}^{7} S_{ic} = 0)$

$D_g = "Undetermined"$

**Else if**$(\sum_{c=1}^{7} S_{ic} = 1)$

Let $\boldsymbol{c}^*$ be a significantly detected category

**If**$(C_{i,c^*} > 0)$ #

$D_g = classfying\ category\ as\ \boldsymbol{c}^*(+)$

**Else**

$D_g = classfying\ category\ as\ \boldsymbol{c}^*(-)$

**Else**

$c^*$ = Max($|C_i|$)

**If**($C_{i,c^*} > 0$)

$D_g = classfying\ category\ as\ \boldsymbol{c}^*(+)$

**Else**

$D_g = classfying\ category\ as\ \boldsymbol{c}^*(-)$

From this decision rule, all genes were classified into 14 categories (7 training patterns (Cluster 1~7) and 7 negative-correlation patterns (Cluster -1~-7)).

***Statistical test for two-group comparison and functional enrichment analysis***

For the gene-set enrichment analysis, DAVID was employed with two biological databases such as gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Huang et al., 2007) using chicken gene IDs. In the DAVID, 5% significance level was considered as biologically significant terms.

***Quantitative real-time PCR (qRT-PCR) analysis of miRNAs***

First-strand cDNA of miRNAs was synthesised from total RNA (1 μg) using the miRNA first-strand cDNA synthesis kit (Stratagene, La Jolla,

CA, USA). To elongate the miRNAs, total RNA was treated with *Escherichia coli* poly(A) polymerase to generate a poly(A) tail at the 3'-end of each RNA molecule. After polyadenylation, cDNAs were synthesised using a real-time adaptor primer. Real-time PCR analysis of the complete miRNA first-strand cDNAs was performed using the High-Specificity miRNA QPCR Core Reagent Kit (Stratagene). The PCR mixture was prepared by adding 2.5 μL 10× core PCR buffer, 2.75 μL 50 mM $MgCl_2$, 10 μL 20 mM deoxynucleotide triphosphate, 1.25 μL 20× EvaGreen, 1.0 μL 3.125 μM universal reverse primer, 1.0 μL 3.125 μM miRNA-specific forward primers, and 0.5 μL high-specificity polymerase to the prepared miRNA cDNA in a 25 μL reaction volume. Each miRNA forward primer and chicken snoRNA primer (Table 4-1) was designed according to the guidelines of Agilent Technologies. miRNA expression was normalised to that of a chicken snoRNA, which has been validated as an internal control (Lee et al., 2011, Rengaraj et al., 2014), and was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

## 3. Results

### *Global expression clustering to understand transcriptional transition among early developmental stage of chicken embryos*

The co-expression analysis was performed to identify the transcriptional transition across all developmental stages, especially during MZT. According to the EGK morphological criteria in the early chicken development, seven representative stages, including oocyte, zygote, and intrauterine embryos (from EGK.I to EGK.X), were employed (Figure 4-2A). Based on the quantification of annotated genes, transcriptional relationships among the samples and stages were examined (Figure 4-1A). In MDS plot, biologically replicated samples were clearly clustered into each developmental stage, as well as distinct dynamic changes in transcriptional features during intrauterine development were found. Next, in the co-expression analysis using WGCNA, eight co-expressed modules were identified (Figure 4-1B). Of eight co-expression patterns, seven representative patterns were determined excluding repetitive patterns (Figure 4-1C), and supervised analysis was performed to detect statistically significant co-expressed genes following such patterns. As a result, diverse gene expression patterns were observed and they are categorized into seven representative patterns considering properties of the correlation measures (Completely opposite correlation patterns were considered as a same category) (Figure 4-2B-D; FDR adjusted $P < 0.1$). In the cluster 1, 703 and 201 positive and negative correlated genes were observed and expression of these genes was changed at two times, between oocyte and zygote, and between EGK.VI and EGK.VIII. Cluster 2, 3, and 7 show first transition between the oocyte and zygote stages. In the cluster 4, only one negative correlated gene was

observed whereas 145 genes were positively correlated with pattern of cluster 4, suggesting only an increasing expression in EGK.VI. Like cluster 4, cluster 5 also showed a transition pattern between EGK.III and EGK.VI, but the genes in the cluster 5 showed a steady increase or decrease in gene expression patterns up to EGK.VIII. In case of cluster 6, 772 and 535 genes were significantly correlated, which shows transition of gene expression between EGK.VI and EGK.VIII. Considering the result of co-expression analysis, it suggests that there are three times of large transcriptional transitions at certain developmental stages such as zygote, EGK.VI, and EGK.VIII during intrauterine development. In particular, the expression levels of a large number of genes were upregulated at zygote and EGK.VI, which suggests that there may be two waves of ZGAs in chicken comparable to other species. Also, we found both increasing and decreasing of gene expression in EGK.VIII.

***Investigating co-expression patterns of genes annotated transcription and translation-related gene ontology (GO) terms***

We expected to find an avian MZT along with the functional terms of transcription and translation, because of coupled ZGA and MZT processes and low content of 28s rRNA of pre-MZT embryo reported in other well-known organisms (Gilbert et al., 2009, Lee et al., 2014). In the co-expression analysis, it was observed that transcriptional changes at zygote and EGK.VI stage. We presumed that at the time point estimated as ZGA, there would be many genes involved in transcription and translation related GO terms. In this regard, we attempted to investigate gene expression patterns of the annotated genes in transcription and translation related GO terms (Figure 4-3A and Table 4-2). As a result, all gene expression levels were co-expressed in their

annotated transcription and translation related GO terms. In the transcription related terms, five terms showed significantly expressed during two stages, zygote (up-regulation of RNA processing and downregulation of transcription repressor activity) and EGK.VI (upregulation of transcription factor activity, regulation of transcription, and transcription). In case of translation related terms, four of the six terms (endoplasmic reticulum, ribosome, structural constituent of ribosome and translation) showed highly down-regulated patterns after fertilization, but all translation-related terms were up-regulated in EGK.VIII. Next, in accordance with the translational terms, changes in the amount of 28s rRNA; the relative amount was reduced markedly after formation of the zygote and recovered gradually after EGK.VIII, as shown in the electrophoresis profiles (Figure 4-3B). These evidences support that MZT occurs in EGK.VIII in the chicken.

***Transition of the molecular functions and signaling pathway during early embryonic development in chicken***

To identify the functional transition of co-expressed genes in specific patterns, an enrichment test was performed based on the GO and KEGG databases (Table 4-3~4-15; enrichment test P-value < 0.05) and the representative terms were summarized (Figure 4-4A). As a result, ion and cation transport were down-regulated from the oocyte to zygote stage; however, they were up-regulated continuously after EGK.VI. Also, three transcription-related terms such as RNA processing, regulation of transcription, positive regulation of transcription, and Basal transcription factors were indicated to be up-regulated after either or both of zygote and EGK.VI (Figure 4-4A). The functional terms such as Ribosome and translation were down-regulated at zygote and upregulated at EGK.VIII

(Figure 4-4A). Additionally, regulation of small GTPase-mediated signal transduction, MAPK signaling pathway, and Notch signaling pathway showed similar transitions at two stages: up-regulated at zygote and down-regulated after EGK.VIII. Similarly, regulation of small GTPase-mediated signal transduction and MAPK signaling pathway by cluster 3 and Notch signaling pathway by cluster 6 (negative correlated) were significantly observed (Figure 4-4B and Table 4-5, 4-14). These indicated that transcriptional and translational dynamics and transition in signaling pathways during MZT occurred in early chicken embryonic development. Wnt signaling was involved in both the transitions, but the Wnt ligand and receptor-related genes were differentially expressed in the two developmental stages (Figure 4-4C). As shown in the heatmap, we observed distinct regulation between the cleavage and area pellucida formation periods. First, *WNT4*, *WNT6*, *FZD1*, *RYK* and *CCNYL1* were up-regulated during cleavage period, from zygote to EGK.VI, and down-regulated after EGK.VI. In contrast, the expression levels of *WNT5B*, *WNT8C*, *DKK1*, *FZD7*, *FZD10* and *CTNNB1* were up-regulated during the area pellucida formation period, from EGK.VI to EGK.VIII.

***Epigenetic dynamics and post-transcriptional regulators in chicken MZT***

Based on the results of the functional analysis, we could speculate global epigenetic status and histone modifications. Genes related to histone H3 acetylation (GO: 0043966) were induced after the zygote stage and decreased after the EGK.VIII stage, which indicates that dynamic modulation occurs in chromatin structures from paternal and maternal pronucleus or zygotic nucleus after fertilization for the ZGA (Figure 4-5A). In addition, gene expression of epigenetic enzymes such as DNA methyltransferases (DNMTs), histone deacetylases (HDACs), methyl binding domains (MBDs),

tet methyl-cytosine dioxygenases (TETs), and histone acetyltransferases (HATs) was investigated during the various developmental stages (Figure 4-5A). In this heatmap, the expression levels of *HDAC4*, *HDAC11*, *MBD3* and *TET2* were down-regulated more than two-fold after the zygote, but these enzymes, except for *TET2*, were up-regulated again after EGK.VIII. Other epigenetic enzymes—*DNMT3A*, *DNMT3B* and *TET1*— were up-regulated in the transition between EGK.VIII and EGK.X, but *DNMT1* and *HDAC9* were down-regulated. These expression patterns seem to be associated with epigenetic reprogramming during MZT in the chicken. To further investigate the roles in maternal RNA clearance of miRNAs in chicken, miRNA machinery genes, such as *DROSHA*, *DICER1*, *DGCR8*, *XPO5* and *AGO2*, were investigated primarily (Figure 4-5B). Of the five representative miRNA machinery examples, *DROSHA* and *DGCR8* were significantly up-regulated during EGK.VIII. These results showed that miRNAs could be actively maturated during EGK.VIII to degrade maternal transcripts through post-translational regulation, because zygote-expressed miRNAs are essential for MZT and maternal mRNA clearance in various species (Svoboda and Flemr, 2010). Generally, strong evidence for MZT can be found in miRNAs sharing the same proximal AAGUGC motif, which has been widely studied in other species, such as *Danio rerio*, *Xenopus laevis* and *Mus musculus* (Svoboda and Flemr, 2010, Yartseva and Giraldez, 2015). Based on this knowledge, a motif search was performed in whole chicken mature miRNAs to detect orthologous miRNAs with the conserved 5'-motif. As a result, the miR-302 family was identified as orthologous to well-known MZT markers, such as dre-miR-430s, xla-miR-427 and mmu-miR-290s (Figure 4-5C). To measure the expression of mature miR-302, qRT-PCR was performed; the results showed significant up-regulation during EGK.VIII (Figure 4-5D and Table 4-16). These results provide strong evidence that MZT occurs in stage EGK.VIII during early

chicken development in the functional aspects.

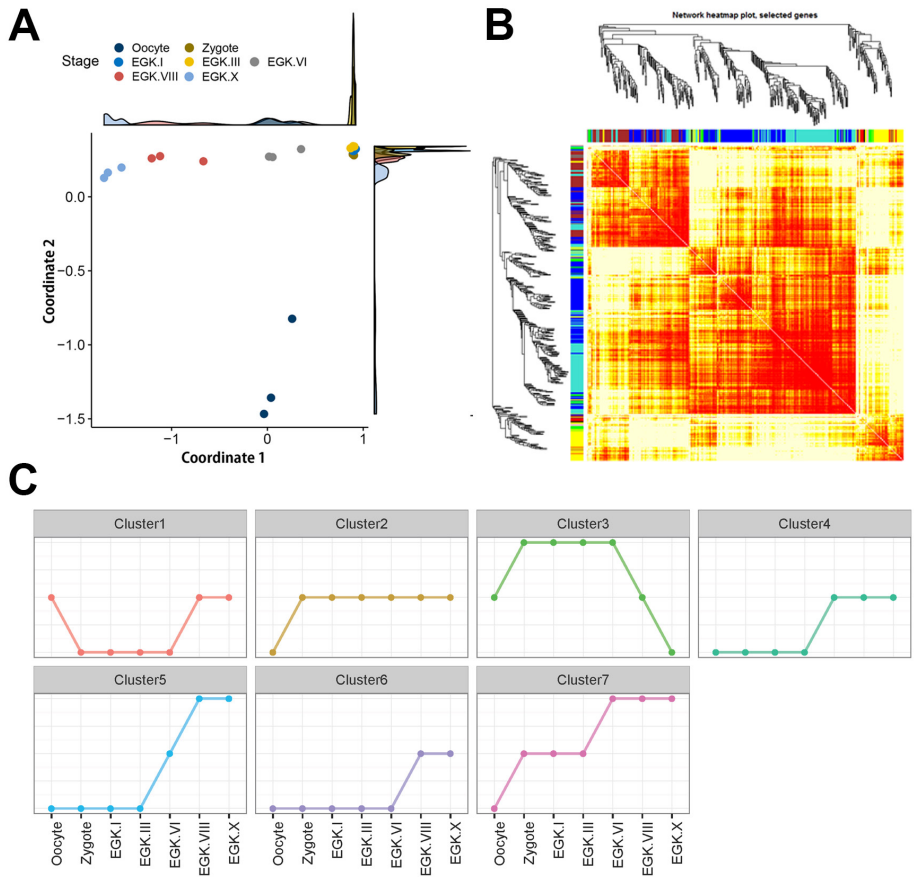


**Figure 4-1. Gene expression clusters of RNA-seq samples.** (A) Multidimensional scaling (MDS) plot, based on gene expression during early development in chicken. Each point represents an RNA-seq sample (B) Result of weighted correlation network analysis (WGCNA). A total of 8 modules were observed in WGCNA analysis. (C) The representative patterns of gene expression in embryogenesis. The clusters show the following expression patterns during early embryogenesis (Cluster 1; degraded at zygote and re-upregulated at EGK.VIII, Cluster 2; induced at zygote and maintained, Cluster 3; induced at zygote and down-regulated from EGK.VI to EGK.X, Cluster 4; up-regulated at EGK.VI, Cluster 5; up-regulated from EGK.III to EGK.VIII, Cluster 6; up-regulated at EGK.VIII, Cluster 7; up-regulated at zygote and EGK.VI).

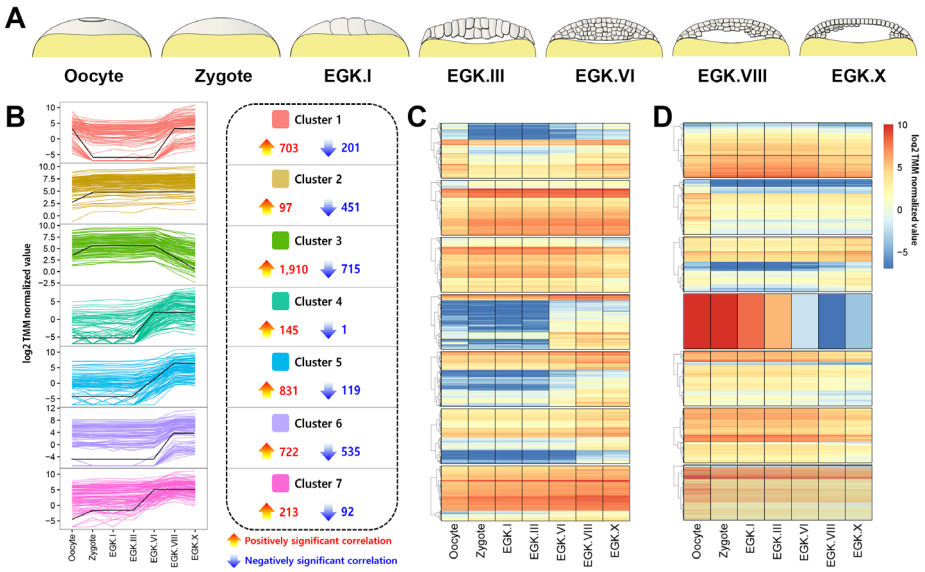


**Figure 4-2. Expression clustering of RNA-seq using early chicken embryos.** (A) Schematic of the morphology of the early chicken embryos from oocyte to EGK.X. (B) Line plots visualizing gene expression patterns across the developmental stages identified in our co-expression analysis. The *y*-axis represents $\log_2$ TMM normalized values, and the bold black line indicates representative patterns in each cluster. All significantly detected positive and negative genes in each cluster are summarized in the right dotted box. (C, D) Heatmap of seven positively (C) and negatively (D) clustered modules from the co-expression analysis. From top to bottom, the heat map shows the significantly enriched genes in clusters, respectively.

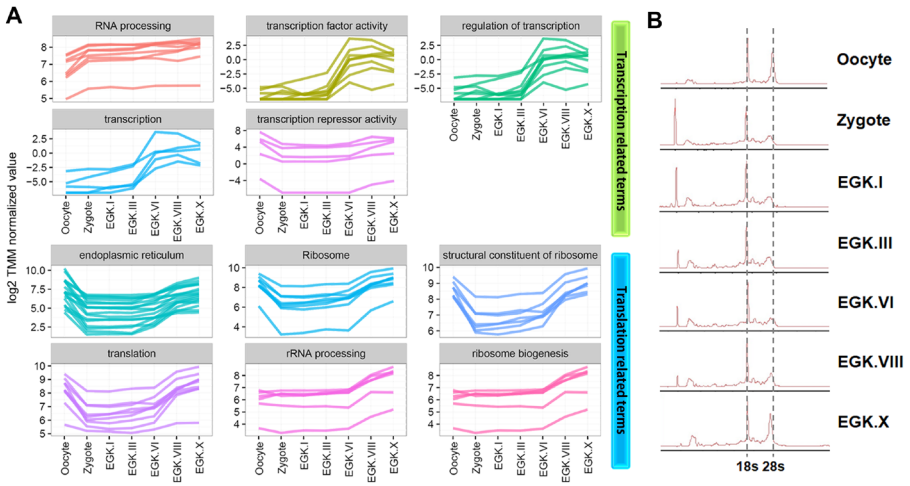


**Figure 4-3. Transcriptional and translational regulation during ZGA and MZT in chicken.** (A) Expression patterns of the clustered genes in relation to biological terms of transcription and translation. (B) Electrophoresis profile of size distribution of the total RNA content during early chicken development. Two dotted lines represent the peak of 18S and 28S rRNA, respectively.

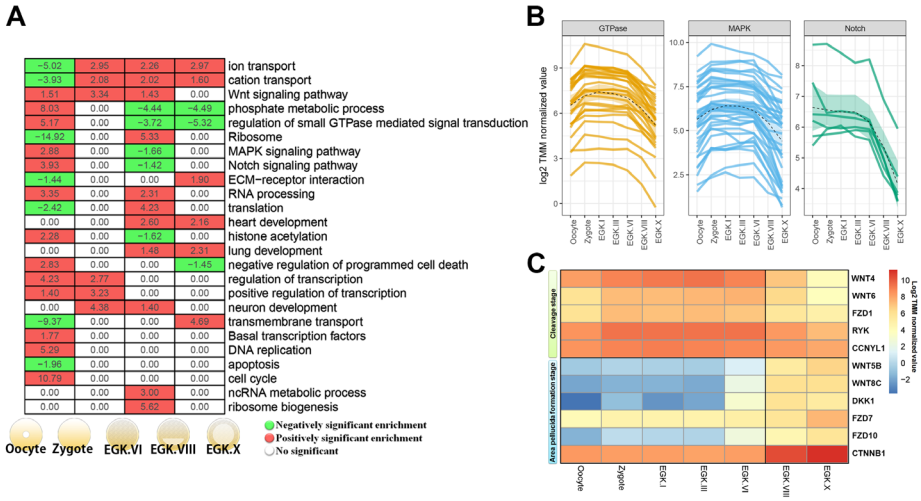


**Figure 4-4. Embryonic developmental transition at EGK.VIII in chicken.** (A) Summary of the significant functional terms across the four transitions. The average $\log_2$ fold-change values of genes belonging to the functional terms are represented in each transition. Positively and negatively enriched terms are filled in green and red, respectively. Non-significantly detected terms during a specific transition are filled in white. (B) The degradation of up-regulated signaling pathways after fertilization from EGK.VIII stage. (C) The detailed gene expression related to Wnt signaling through early embryogenesis.

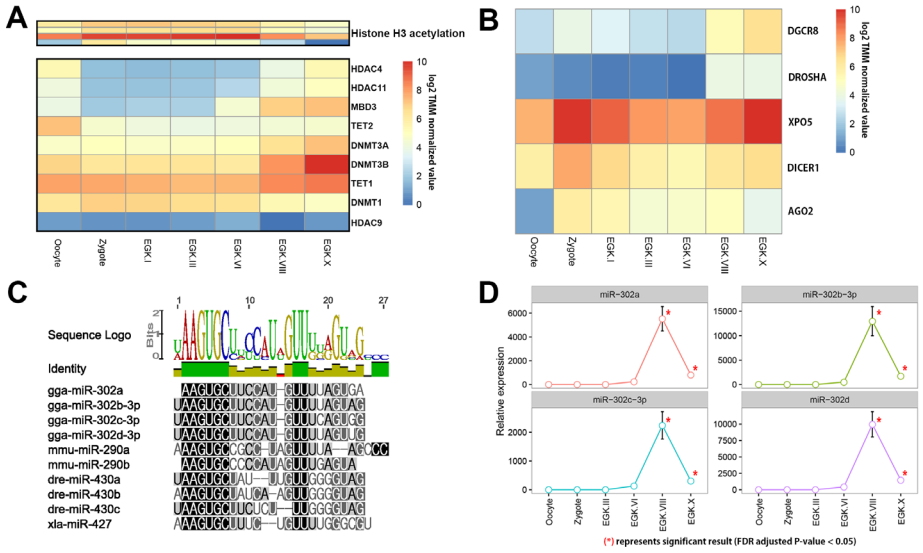


**Figure 4-5. Epigenetic dynamics in early chicken development.** (A) The expression patterns of Histone H3 acetylation and epigenetic enzymes. (B) The expression profiling of miRNA machinery genes including DGCR8, DROSHA, XPO5, DICER1, and AGO2 through early embryogenesis. (C) Multiple sequence alignment of several mature miRNAs. From 2 to 7 bp, the 5'-proximal AAGUGC motif was conserved. (D) Mature gga-miR-302 family expression patterns in early chicken development. qRT-PCR was conducted to quantify gene expression, and relative gene expression was measured by comparison with the control gene encoding a snoRNA. Error bars indicate the SE among three biological replicates.

**Table 4-1. Primers used for quantitative real-time PCR analysis of miRNAs**

| miRNA | miRNA sequence | Forward primer sequence (5'→3') |
|---|---|---|
| gga-miR-302a | AAGUGCUUCCAUGUUUUAGUGA | GAAGTGCTTCCATGTTTTAGTGA |
| gga-miR-302b-3p | UAAGUGCUUCCAUGUUUUAGUAG | GGTAAGTGCTTCCATGTTTTAGTAG |
| gga-miR-302c-3p | UAAGUGCUUCCAUGUUUCAGUGG | TAAGTGCTTCCATGTTTCAGTGG |
| gga-miR-302d-3p | UAAGUGCUUCCAUGUUUUAGUUG | GTAAGTGCTTCCATGTTTTAGTTG |
| snoRNA | | GGGATGTAAAAAAATACTTGCTATC |

**Table 4-2. Transcription and translation related terms and allocated genes in those terms**

| Terms | Genes | P-value | Cluster | Type |
|---|---|---|---|---|
| RNA processing | ENSGALG00000008372<br>ENSGALG00000013064<br>ENSGALG00000010651<br>ENSGALG00000015712<br>ENSGALG00000004555<br>ENSGALG00000010501<br>ENSGALG00000003906<br>ENSGALG00000009001 | 0.037157 | Cluster 2 | Transcription |
| transcription factor activity | ENSGALG00000006365<br>ENSGALG00000009274<br>ENSGALG00000010436<br>ENSGALG00000000365<br>ENSGALG00000019842<br>ENSGALG00000009277<br>ENSGALG00000008848<br>ENSGALG00000011053 | 0.001383 | Cluster 4 | Transcription |
| regulation of transcription | ENSGALG00000006365<br>ENSGALG00000009274<br>ENSGALG00000010436<br>ENSGALG00000000365<br>ENSGALG00000023296<br>ENSGALG00000019842<br>ENSGALG00000009277<br>ENSGALG00000008848<br>ENSGALG00000011053 | 0.005168 | Cluster 4 | Transcription |
| transcription | ENSGALG00000006365<br>ENSGALG00000009274<br>ENSGALG00000023296<br>ENSGALG00000009277<br>ENSGALG00000008848 | 0.032068 | Cluster 4 | Transcription |
| transcription repressor activity | ENSGALG00000015184<br>ENSGALG00000007357<br>ENSGALG00000004274<br>ENSGALG00000003272<br>ENSGALG00000004288 | 0.059575 | Cluster 1 | Transcription |
| endoplasmic reticulum | ENSGALG00000007233<br>ENSGALG00000008000<br>ENSGALG00000002844<br>ENSGALG00000016386<br>ENSGALG00000012726<br>ENSGALG00000009330<br>ENSGALG00000015612<br>ENSGALG00000010073<br>ENSGALG00000013462<br>ENSGALG00000008676<br>ENSGALG00000008348<br>ENSGALG00000010991<br>ENSGALG00000016909 | 5.71E-04 | Cluster 1 | Translation |

| | | | | |
|---|---|---|---|---|
| | ENSGALG00000019891<br>ENSGALG00000005945<br>ENSGALG00000001053<br>ENSGALG00000004833 | | | |
| Ribosome | ENSGALG00000001658<br>ENSGALG00000007611<br>ENSGALG00000005338<br>ENSGALG00000015195<br>ENSGALG00000014833<br>ENSGALG00000011290<br>ENSGALG00000006096<br>ENSGALG00000002644 | 0.00543<br>3 | Cluster<br>1 | Translation |
| structural constituent of ribosome | ENSGALG00000001658<br>ENSGALG00000007611<br>ENSGALG00000015195<br>ENSGALG00000014833<br>ENSGALG00000011290<br>ENSGALG00000006096<br>ENSGALG00000002644 | 0.02842<br>9 | Cluster<br>1 | Translation |
| translation | ENSGALG00000001658<br>ENSGALG00000007611<br>ENSGALG00000015195<br>ENSGALG00000000907<br>ENSGALG00000014833<br>ENSGALG00000011290<br>ENSGALG00000011941<br>ENSGALG00000006096<br>ENSGALG00000002644 | 0.08948<br>7 | Cluster<br>1 | Translation |
| rRNA processing | ENSGALG00000012916<br>ENSGALG00000014791<br>ENSGALG00000000247<br>ENSGALG00000012636<br>ENSGALG00000016073<br>ENSGALG00000005558 | 9.19E-<br>04 | Cluster<br>6 | Translation |
| ribosome biogenesis | ENSGALG00000012916<br>ENSGALG00000014791<br>ENSGALG00000000247<br>ENSGALG00000012636<br>ENSGALG00000016073<br>ENSGALG00000005558 | 0.00435<br>7 | Cluster<br>6 | Translation |

## Table 4-3. Significantly detected GO-BP and KEGG terms for cluster 1

**Significantly detected GO-BP terms for cluster 1 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| lysosome organization | 4 | 0.7 | 1.70E-03 |
| cell redox homeostasis | 6 | 1.1 | 5.80E-03 |
| vacuole organization | 4 | 0.7 | 6.00E-03 |
| associative learning | 3 | 0.5 | 2.00E-02 |
| phosphorylation | 23 | 4 | 3.00E-02 |
| translation | 12 | 2.1 | 3.00E-02 |
| protein amino acid phosphorylation | 20 | 3.5 | 3.90E-02 |
| negative regulation of cell adhesion | 3 | 0.5 | 4.00E-02 |

**Significantly detected KEGG pathways for cluster 1 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Lysosome | 13 | 2.3 | 1.50E-04 |
| Ribosome | 11 | 1.9 | 2.90E-04 |
| Glycosaminoglycan degradation | 5 | 0.9 | 4.80E-03 |

## Table 4-4. Significantly detected GO-BP and KEGG terms for cluster 2

**Significantly detected GO-BP terms for cluster 2 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| cellular macromolecule catabolic process | 5 | 5.6 | 1.40E-02 |
| M phase of mitotic cell cycle | 3 | 3.4 | 1.90E-02 |
| macromolecule catabolic process | 5 | 5.6 | 2.10E-02 |
| cell cycle process | 4 | 4.5 | 2.20E-02 |
| RNA methylation | 2 | 2.2 | 3.90E-02 |
| M phase | 3 | 3.4 | 4.10E-02 |
| mitotic cell cycle | 3 | 3.4 | 4.30E-02 |
| RNA processing | 4 | 4.5 | 4.40E-02 |

**Significantly detected KEGG pathways for cluster 2 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Cell cycle | 5 | 5.6 | 2.20E-03 |

## Table 4-5. Significantly detected GO-BP and KEGG terms for cluster 3

**Significantly detected GO-BP terms for cluster 3 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| phosphorus metabolic process | 89 | 5.2 | 4.70E-08 |
| phosphate metabolic process | 89 | 5.2 | 4.70E-08 |
| protein amino acid phosphorylation | 70 | 4.1 | 5.90E-08 |
| phosphorylation | 76 | 4.4 | 3.30E-07 |
| regulation of small GTPase mediated signal transduction | 28 | 1.6 | 5.90E-05 |
| regulation of Ras protein signal transduction | 24 | 1.4 | 8.30E-05 |
| regulation of Rab protein signal transduction | 11 | 0.6 | 3.10E-04 |
| regulation of Rab GTPase activity | 11 | 0.6 | 3.10E-04 |
| protein kinase cascade | 18 | 1.1 | 4.10E-04 |
| intracellular signaling cascade | 56 | 3.3 | 8.80E-04 |
| regulation of Ras GTPase activity | 13 | 0.8 | 9.90E-04 |
| response to DNA damage stimulus | 25 | 1.5 | 1.10E-03 |
| regulation of GTPase activity | 13 | 0.8 | 2.60E-03 |
| cellular response to stress | 28 | 1.6 | 3.40E-03 |
| myeloid leukocyte activation | 6 | 0.4 | 3.60E-03 |
| DNA repair | 20 | 1.2 | 4.20E-03 |
| regulation of apoptosis | 31 | 1.8 | 1.50E-02 |
| regulation of programmed cell death | 31 | 1.8 | 2.00E-02 |
| chromatin modification | 14 | 0.8 | 2.00E-02 |
| sexual reproduction | 15 | 0.9 | 2.10E-02 |
| regulation of protein kinase activity | 15 | 0.9 | 2.10E-02 |
| regulation of cell death | 31 | 1.8 | 2.10E-02 |
| histone modification | 9 | 0.5 | 2.20E-02 |
| vacuolar transport | 4 | 0.2 | 2.30E-02 |
| regulation of kinase activity | 15 | 0.9 | 2.50E-02 |
| regulation of hydrolase activity | 16 | 0.9 | 2.80E-02 |
| covalent chromatin modification | 9 | 0.5 | 2.90E-02 |
| DNA metabolic process | 31 | 1.8 | 2.90E-02 |
| regulation of transferase activity | 15 | 0.9 | 3.00E-02 |
| positive regulation of organelle organization | 7 | 0.4 | 3.30E-02 |
| chromatin organization | 19 | 1.1 | 3.40E-02 |
| keratinocyte differentiation | 4 | 0.2 | 3.40E-02 |
| histone H3 acetylation | 4 | 0.2 | 3.40E-02 |
| regulation of interleukin-2 production | 4 | 0.2 | 3.40E-02 |
| programmed cell death | 17 | 1 | 3.80E-02 |
| negative regulation of apoptosis | 17 | 1 | 4.10E-02 |
| regulation of cell cycle | 14 | 0.8 | 4.30E-02 |
| regulation of cellular response to stress | 6 | 0.4 | 4.40E-02 |
| negative regulation of cell death | 17 | 1 | 4.80E-02 |
| negative regulation of programmed cell death | 17 | 1 | 4.80E-02 |
| endosome to lysosome transport | 3 | 0.2 | 4.80E-02 |
| macrophage activation | 3 | 0.2 | 4.80E-02 |
| anti-apoptosis | 9 | 0.5 | 4.90E-02 |

**Significantly detected KEGG pathways for cluster 3 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| MAPK signaling pathway | 37 | 2.2 | 1.60E-03 |
| Endocytosis | 29 | 1.7 | 5.00E-03 |
| Glycerolipid metabolism | 10 | 0.6 | 1.60E-02 |
| ErbB signaling pathway | 15 | 0.9 | 2.60E-02 |
| Progesterone-mediated oocyte maturation | 14 | 0.8 | 2.90E-02 |
| VEGF signaling pathway | 12 | 0.7 | 4.40E-02 |

## Table 4-6. Significantly detected GO-BP terms for cluster 4

**Significantly detected GO-BP terms for cluster 4 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| embryonic morphogenesis | 7 | 8 | 3.50E-05 |
| regulation of transcription, DNA-dependent | 11 | 12.5 | 2.20E-04 |
| regulation of RNA metabolic process | 11 | 12.5 | 2.50E-04 |
| embryonic organ morphogenesis | 5 | 5.7 | 2.70E-04 |
| regulation of transcription | 12 | 13.6 | 6.10E-04 |
| embryonic organ development | 5 | 5.7 | 7.40E-04 |
| sensory organ development | 5 | 5.7 | 1.00E-03 |
| ear development | 4 | 4.5 | 1.20E-03 |
| skeletal system development | 5 | 5.7 | 1.90E-03 |
| inner ear morphogenesis | 3 | 3.4 | 1.10E-02 |
| ear morphogenesis | 3 | 3.4 | 1.20E-02 |
| regulation of transcription from RNA polymerase II promoter | 5 | 5.7 | 1.30E-02 |
| detection of mechanical stimulus involved in sensory perception of sound | 2 | 2.3 | 1.30E-02 |
| inner ear development | 3 | 3.4 | 1.60E-02 |
| detection of mechanical stimulus involved in sensory perception | 2 | 2.3 | 1.80E-02 |
| pattern specification process | 4 | 4.5 | 1.90E-02 |
| transcription | 6 | 6.8 | 2.00E-02 |
| detection of mechanical stimulus | 2 | 2.3 | 2.20E-02 |
| limb morphogenesis | 3 | 3.4 | 2.40E-02 |
| appendage morphogenesis | 3 | 3.4 | 2.40E-02 |
| limb development | 3 | 3.4 | 2.70E-02 |
| appendage development | 3 | 3.4 | 2.70E-02 |
| retina morphogenesis in camera-type eye | 2 | 2.3 | 2.70E-02 |
| gland development | 3 | 3.4 | 3.00E-02 |
| detection of stimulus involved in sensory perception | 2 | 2.3 | 3.10E-02 |
| forebrain development | 3 | 3.4 | 3.40E-02 |
| male genitalia development | 2 | 2.3 | 3.50E-02 |
| response to mechanical stimulus | 2 | 2.3 | 3.50E-02 |
| cell fate commitment | 3 | 3.4 | 4.00E-02 |
| negative regulation of Wnt receptor signaling pathway | 2 | 2.3 | 4.40E-02 |

## Table 4-7. Significantly detected GO-BP and KEGG terms for cluster 5

**Significantly detected GO-BP terms for cluster 5 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| potassium ion transport | 12 | 1.7 | 4.00E-03 |
| cellular component morphogenesis | 14 | 2 | 4.90E-03 |
| gastrulation | 7 | 1 | 5.40E-03 |
| cell projection morphogenesis | 10 | 1.4 | 5.40E-03 |
| transmembrane receptor protein tyrosine kinase signaling pathway | 12 | 1.7 | 5.70E-03 |
| neuron differentiation | 15 | 2.1 | 5.90E-03 |
| placenta development | 6 | 0.9 | 6.00E-03 |
| neuron development | 12 | 1.7 | 7.20E-03 |
| segmentation | 6 | 0.9 | 7.20E-03 |
| ribosome biogenesis | 7 | 1 | 7.30E-03 |
| neuron projection development | 10 | 1.4 | 8.00E-03 |
| axonogenesis | 9 | 1.3 | 8.00E-03 |
| cell part morphogenesis | 10 | 1.4 | 8.80E-03 |
| ncRNA metabolic process | 11 | 1.6 | 9.10E-03 |
| neuron projection morphogenesis | 9 | 1.3 | 9.80E-03 |
| cell morphogenesis | 12 | 1.7 | 1.10E-02 |
| ion transport | 31 | 4.4 | 1.10E-02 |
| nuclear export | 4 | 0.6 | 1.20E-02 |
| hippocampus development | 4 | 0.6 | 1.20E-02 |
| cellular amino acid derivative biosynthetic process | 6 | 0.9 | 1.20E-02 |
| metal ion transport | 19 | 2.7 | 1.20E-02 |
| axon guidance | 7 | 1 | 1.20E-02 |
| pattern specification process | 14 | 2 | 1.30E-02 |
| cell morphogenesis involved in differentiation | 10 | 1.4 | 1.30E-02 |
| cell morphogenesis involved in neuron differentiation | 9 | 1.3 | 1.40E-02 |
| endocrine system development | 6 | 0.9 | 1.60E-02 |
| monovalent inorganic cation transport | 16 | 2.3 | 1.60E-02 |
| determination of left/right symmetry | 5 | 0.7 | 1.60E-02 |
| biogenic amine biosynthetic process | 5 | 0.7 | 1.60E-02 |
| determination of symmetry | 5 | 0.7 | 1.90E-02 |
| determination of bilateral symmetry | 5 | 0.7 | 1.90E-02 |
| somitogenesis | 5 | 0.7 | 1.90E-02 |
| embryonic organ development | 11 | 1.6 | 1.90E-02 |
| RNA processing | 14 | 2 | 2.20E-02 |
| RNA export from nucleus | 3 | 0.4 | 2.30E-02 |
| limbic system development | 4 | 0.6 | 2.50E-02 |
| regionalization | 11 | 1.6 | 2.50E-02 |
| ribonucleoprotein complex biogenesis | 7 | 1 | 2.70E-02 |
| protein targeting | 7 | 1 | 2.70E-02 |
| anterior/posterior pattern formation | 9 | 1.3 | 2.70E-02 |
| cell projection organization | 11 | 1.6 | 3.00E-02 |
| embryonic morphogenesis | 14 | 2 | 3.20E-02 |
| rRNA metabolic process | 5 | 0.7 | 3.50E-02 |
| rRNA processing | 5 | 0.7 | 3.50E-02 |
| RNA transport | 4 | 0.6 | 3.70E-02 |

| | | | |
|---|---|---|---|
| establishment of RNA localization | 4 | 0.6 | 3.70E-02 |
| nucleobase, nucleoside, nucleotide and nucleic acid transport | 4 | 0.6 | 3.70E-02 |
| nucleic acid transport | 4 | 0.6 | 3.70E-02 |
| enzyme linked receptor protein signaling pathway | 13 | 1.9 | 3.70E-02 |
| tube development | 11 | 1.6 | 3.80E-02 |
| mesoderm development | 5 | 0.7 | 3.90E-02 |
| RNA localization | 4 | 0.6 | 4.30E-02 |
| forebrain development | 8 | 1.1 | 4.40E-02 |
| telencephalon development | 5 | 0.7 | 5.00E-02 |

**Significantly detected KEGG pathways for cluster 5 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Spliceosome | 13 | 1.9 | 1.20E-03 |
| Melanogenesis | 11 | 1.6 | 2.90E-03 |
| Calcium signaling pathway | 15 | 2.1 | 3.20E-03 |
| Neuroactive ligand-receptor interaction | 20 | 2.9 | 3.70E-03 |
| Wnt signaling pathway | 12 | 1.7 | 1.90E-02 |

## Table 4-8. Significantly detected GO-BP and KEGG terms for cluster 6

**Significantly detected GO-BP terms for cluster 6 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| ribosome biogenesis | 7 | 1.1 | 1.60E-03 |
| rRNA processing | 6 | 1 | 1.90E-03 |
| rRNA metabolic process | 6 | 1 | 1.90E-03 |
| ribonucleoprotein complex biogenesis | 7 | 1.1 | 6.40E-03 |
| glutamine family amino acid metabolic process | 5 | 0.8 | 1.10E-02 |
| ncRNA metabolic process | 9 | 1.5 | 1.20E-02 |
| RNA processing | 12 | 2 | 1.30E-02 |
| ncRNA processing | 7 | 1.1 | 1.50E-02 |
| in utero embryonic development | 8 | 1.3 | 2.10E-02 |
| Wnt receptor signaling pathway | 7 | 1.1 | 2.20E-02 |
| positive regulation of biosynthetic process | 16 | 2.6 | 2.50E-02 |
| positive regulation of macromolecule biosynthetic process | 15 | 2.4 | 3.80E-02 |
| protein amino acid autophosphorylation | 4 | 0.7 | 4.00E-02 |
| positive regulation of cellular biosynthetic process | 15 | 2.4 | 4.70E-02 |
| regulation of blood coagulation | 3 | 0.5 | 5.00E-02 |
| glutamine metabolic process | 3 | 0.5 | 5.00E-02 |

**Significantly detected KEGG pathways for cluster 6 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Alanine, aspartate and glutamate metabolism | 5 | 0.8 | 1.80E-02 |
| Neuroactive ligand-receptor interaction | 16 | 2.6 | 4.70E-02 |

## Table 4-9. Significantly detected GO-BP and KEGG terms for cluster 7

**Significantly detected GO-BP terms for cluster 7 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| cell division | 6 | 3.2 | 8.40E-04 |
| mitotic cell cycle | 6 | 3.2 | 1.10E-03 |
| nuclear division | 5 | 2.7 | 1.40E-03 |
| mitosis | 5 | 2.7 | 1.40E-03 |
| M phase of mitotic cell cycle | 5 | 2.7 | 1.70E-03 |
| organelle fission | 5 | 2.7 | 1.80E-03 |
| translation | 9 | 4.9 | 2.70E-03 |
| cell cycle phase | 6 | 3.2 | 3.60E-03 |
| cell cycle | 8 | 4.3 | 3.80E-03 |
| M phase | 5 | 2.7 | 7.30E-03 |
| DNA integrity checkpoint | 3 | 1.6 | 9.30E-03 |
| DNA damage checkpoint | 3 | 1.6 | 9.30E-03 |
| RNA processing | 7 | 3.8 | 1.10E-02 |
| cell cycle process | 6 | 3.2 | 1.40E-02 |
| cell cycle checkpoint | 3 | 1.6 | 1.50E-02 |
| RNA splicing | 4 | 2.2 | 1.80E-02 |
| chromosome segregation | 3 | 1.6 | 2.20E-02 |
| transcription | 11 | 5.9 | 2.30E-02 |
| DNA damage response, signal transduction | 3 | 1.6 | 2.70E-02 |
| response to DNA damage stimulus | 6 | 3.2 | 3.10E-02 |
| macromolecular complex subunit organization | 6 | 3.2 | 4.90E-02 |

**Significantly detected KEGG pathways for cluster 7 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Spliceosome | 11 | 5.9 | 5.40E-07 |
| Cell cycle | 7 | 3.8 | 2.60E-03 |
| Pyrimidine metabolism | 5 | 2.7 | 2.10E-02 |
| RNA degradation | 4 | 2.2 | 2.40E-02 |
| Basal transcription factors | 3 | 1.6 | 4.40E-02 |

## Table 4-10. Significantly detected GO-BP and KEGG terms for cluster -1

**Significantly detected GO-BP terms for cluster -1 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| cellular response to stress | 9 | 4.8 | 2.42E-03 |
| response to DNA damage stimulus | 8 | 4.3 | 2.51E-03 |
| DNA metabolic process | 9 | 4.8 | 1.38E-02 |
| DNA repair | 6 | 3.2 | 1.70E-02 |
| DNA replication | 5 | 2.7 | 4.24E-02 |

## Table 4-11. Significantly detected GO-BP and KEGG terms for cluster -2

**Significantly detected GO-BP terms for cluster -2 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| steroid biosynthetic process | 6 | 1.5 | 2.40E-04 |
| cell-matrix adhesion | 5 | 1.3 | 4.75E-03 |
| cell-substrate adhesion | 5 | 1.3 | 6.24E-03 |
| aminoglycan metabolic process | 4 | 1.0 | 6.85E-03 |
| enzyme linked receptor protein signaling pathway | 10 | 2.6 | 9.75E-03 |
| cholesterol biosynthetic process | 3 | 0.8 | 1.00E-02 |
| transmembrane receptor protein tyrosine kinase signaling pathway | 8 | 2.1 | 1.09E-02 |
| steroid metabolic process | 6 | 1.5 | 1.34E-02 |
| transmembrane transport | 17 | 4.4 | 1.50E-02 |
| interphase of mitotic cell cycle | 4 | 1.0 | 1.56E-02 |
| interphase | 4 | 1.0 | 1.56E-02 |
| sterol biosynthetic process | 3 | 0.8 | 1.81E-02 |
| response to peptide hormone stimulus | 4 | 1.0 | 2.28E-02 |
| cellular response to hormone stimulus | 4 | 1.0 | 2.85E-02 |
| lipid localization | 5 | 1.3 | 3.20E-02 |
| G1/S transition of mitotic cell cycle | 3 | 0.8 | 3.37E-02 |
| polysaccharide metabolic process | 4 | 1.0 | 4.56E-02 |
| cell adhesion | 13 | 3.4 | 4.67E-02 |
| biological adhesion | 13 | 3.4 | 4.67E-02 |

**Significantly detected KEGG pathways for cluster -2 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Lysosome | 9 | 2.3 | 4.13E-03 |
| Butanoate metabolism | 5 | 1.3 | 4.21E-03 |
| Other glycan degradation | 4 | 1.0 | 4.62E-03 |
| Steroid biosynthesis | 4 | 1.0 | 4.62E-03 |
| Valine, leucine and isoleucine degradation | 5 | 1.3 | 2.09E-02 |
| Glycosphingolipid biosynthesis | 3 | 0.8 | 4.76E-02 |

## Table 4-12. Significantly detected GO-BP and KEGG terms for cluster -3

**Significantly detected GO-BP terms for cluster -3 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| translation | 17 | 2.7 | 2.93E-03 |
| regulation of filopodium assembly | 3 | 0.5 | 5.06E-03 |
| positive regulation of filopodium assembly | 3 | 0.5 | 5.06E-03 |
| regulation of cell projection assembly | 3 | 0.5 | 9.85E-03 |
| positive regulation of cellular component organization | 7 | 1.1 | 1.15E-02 |
| transmembrane transport | 23 | 3.7 | 1.82E-02 |
| regulation of establishment of protein localization | 5 | 0.8 | 1.88E-02 |
| skeletal system development | 11 | 1.7 | 2.91E-02 |
| regulation of protein localization | 5 | 0.8 | 3.10E-02 |
| tRNA modification | 3 | 0.5 | 3.17E-02 |
| regulation of cellular component biogenesis | 6 | 1.0 | 3.81E-02 |
| heart development | 9 | 1.4 | 4.63E-02 |

**Significantly detected KEGG pathways for cluster -3 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Proteasome | 9 | 1.4 | 1.06E-04 |

## Table 4-13. Significantly detected GO-BP and KEGG terms for cluster -5

**Significantly detected GO-BP terms for cluster -5 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| cell projection assembly | 3 | 2.8 | 9.00E-03 |
| regulation of Ras GTPase activity | 3 | 2.8 | 2.33E-02 |
| regulation of GTPase activity | 3 | 2.8 | 2.82E-02 |
| filopodium assembly | 2 | 1.9 | 3.07E-02 |
| vesicle-mediated transport | 4 | 3.8 | 3.54E-02 |
| microspike assembly | 2 | 1.9 | 3.57E-02 |

**Significantly detected KEGG pathways for cluster -5 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Focal adhesion | 5 | 4.7 | 2.13E-02 |
| Inositol phosphate metabolism | 3 | 2.8 | 3.43E-02 |

## Table 4-14. Significantly detected GO-BP and KEGG terms for cluster -6

**Significantly detected GO-BP terms for cluster -6 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| autophagy | 5 | 1.0 | 2.80E-05 |
| protein localization | 17 | 3.4 | 5.69E-03 |
| protein transport | 15 | 3.0 | 8.62E-03 |
| establishment of protein localization | 15 | 3.0 | 8.62E-03 |
| regulation of small GTPase mediated signal transduction | 10 | 2.0 | 1.65E-02 |
| response to protein stimulus | 4 | 0.8 | 2.52E-02 |
| cellular protein localization | 10 | 2.0 | 2.55E-02 |
| cellular macromolecule localization | 10 | 2.0 | 2.77E-02 |
| intracellular transport | 12 | 2.4 | 2.93E-02 |
| dephosphorylation | 7 | 1.4 | 3.74E-02 |
| intracellular protein transport | 9 | 1.8 | 3.77E-02 |
| regulation of Ras protein signal transduction | 8 | 1.6 | 3.83E-02 |
| regulation of cell growth | 5 | 1.0 | 4.49E-02 |

**Significantly detected KEGG pathways for cluster -6 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Valine, leucine and isoleucine degradation | 8 | 1.6 | 3.59E-04 |
| Notch signaling pathway | 6 | 1.2 | 1.96E-02 |
| Regulation of autophagy | 4 | 0.8 | 2.59E-02 |
| SNARE interactions in vesicular transport | 5 | 1.0 | 2.99E-02 |

## Table 4-15. Significantly detected GO-BP and KEGG terms for cluster -7

**Significantly detected GO-BP terms for cluster -7 (enrichment test P-value < 0.05)**

| GO term | Count | % | P-Value |
|---|---|---|---|
| transmembrane receptor protein serine/threonine kinase signaling pathway | 3 | 3.6 | 2.14E-02 |
| calcium ion transport | 3 | 3.6 | 4.13E-02 |
| di-, tri-valent inorganic cation transport | 3 | 3.6 | 4.95E-02 |

**Significantly detected KEGG pathways for cluster -7 (enrichment test P-value < 0.05)**

| KEGG term | Count | % | P-Value |
|---|---|---|---|
| Lysosome | 4 | 4.8 | 2.85E-02 |

**Table 4-16. Result of statistical analysis for miR-302 family**

| miR-302 family | Oocyte vs Zygote | Zygote vs EGK.III | EGK.III vs EGK.VI | EGK.VI vs EGK.VIII | EGK.VIII vs EGK.X |
|---|---|---|---|---|---|
| miR-302a | 0.999939 | 0.999939 | 0.870513 | 5.16E-06 | 1.32E-05 |
| miR-302b-3p | 0.999879 | 0.999879 | 0.970744 | 4.25E-05 | 9.02E-05 |
| miR-302c-3p | 0.999925 | 0.999925 | 0.810899 | 2.41E-05 | 4.53E-05 |
| miR-302d | 0.999979 | 0.999979 | 0.889321 | 7.78E-06 | 1.99E-05 |

These values represents FDR adjusted P-values.

## 4. Discussion

During MZT, the maternal program is cleared and the zygotic developmental program is induced in vertebrates. Thus, this period is the crucial point of becoming independent organism with respect to morphological and genetic changes. In birds, the MBT seems to occur at the beginning of second half of intrauterine development, area pellucida formation around EGK.VI to EGK.VIII (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980, Park et al., 2006). In addition to morphological concept, we examined the systematic transition of transcriptional and translational dynamics based on RNA-seq in chicken, a representative model of avian species.

Our transcriptomic and co-expression patterned analysis indicated that two times of transcriptional transition between oocyte and zygote, and EGK.III and EGK.VI. In mammals, the first transcriptional activation, so called minor ZGA, occurs during the pronucelus stage due to the reprogramming of chromatin state and histone modifications (Bouniol et al., 1995, Aoki et al., 1997, Aoshima et al., 2015, Wu et al., 2016, Dahl et al., 2016). Similarly, after fertilization in the chicken, certain gene sets were affected, including up-regulation of histone H3 acetylation and down-regulation of several epigenetic modulators involved in transcriptional repression: *HDAC4*, *HDAC11* and *MBD3*.

After the zygote upon fertilization, no dynamic changes in transcription were found until EGK.VI based on gene expression clusters. During these cleavage stages, a rapid celluarization takes place and MBT-like event such as cell cycle lengthen appear around EGK.VI based on cell counting analysis (Park et al., 2006, Nagai et al., 2015). Coincidently, the

second transcriptional activation in little later point between EGK.III and EGK.VI were observed although the phosphorylation of RNA polymerase II C-terminal domain (p-PolII) started to be apparent during the late EGK.II to early EGK.III (Nagai et al., 2015).

From the transitional point of EGK.VI, the definite morphological and genetic changes including the establishment of naïve pluripotent state are progressed (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980, Mak et al., 2015). In addition to transcriptional dynamics, functional and translational components clearly suggest that EGK.VIII stage is critical point for MZT in chicken. Of significant functional terms, the gene sets such as regulation of small GTPase-mediated signal transduction, MAPK signaling pathway, and Notch signaling pathway were firstly down-regulated between EGK.VI and EGK.VIII, suggesting the clearance of maternal mRNA. Interestingly, Wnt signaling pathway is up-regulated in both transitions and the associated genes including ligands and receptors were converted after MZT. In a previous report, pre-MZT bovine embryos showed atypically low levels of 28S rRNA (Gilbert et al., 2009); this was also observed in our study, from the zygote to EGK.VI stages, in the chicken. We observed that the 28S rRNA level recovered following MZT, which follows EGK.VIII. These findings were consistent with the translation-related genes significantly down-regulated during the zygote stage and up-regulated during the EGK.VIII stage.

Indeed, during those stages, epigenetic modulators such as *DNMT3A*, *DNMT3B* and *TET1*, which regulate DNA methylation and de-methylation, were up-regulated again during EGK.VIII. These findings are also coincident with the global reduction in transcriptional activity during EGK.VIII (Lee et al., 2016), which seems to represent the epigenetic reprogramming along with

MZT at the cellular level, as in the case of *Xenopus* and zebrafish (Akkers et al., 2009, Lindeman et al., 2011). Moreover, zygotic expression of miRNAs is essential for MZT and maternal mRNA degradation. Based on our expression analysis of miRNA machinery, along with previous reports which showed that the miR-302 family was abundantly expressed in chicken EGK.X blastoderm (Lee et al., 2011) as was the miR-290 family in the mouse embryo after ZGA (Svoboda and Flemr, 2010), we confirmed that the expression of mature miR-302 members containing the proximal AAGUGC motif increased during EGK.VIII. Previous reports. Thus, the miR-302 family may regulate MZT and maternal mRNA degradation.

In conclusion, we demonstrated the transcriptomic transitions of early developmental stages and the first establishment of embryonic identity during MZT in chicken, which was previously limited to morphological studies. Taking all evidence together, we conclude that MZT proceeds from EGK.VIII stage after transcriptional activation in EGK.VI stage during early chicken development. Our results open insights into the issues of MZT mechanism among vertebrates.

# CHAPTER 5

# A Series of Transcription Factors Identified during Early Chicken Embryogenesis for Application to Reprogramming in Birds

## 1. Introduction

In vertebrates, the totipotent zygote initiates well-organized genetic programs establishing of the individual organism upon fertilization. The first round of early development with daughter cell division is regulated by maternal factors in the oocyte (Tadros and Lipshitz, 2009), until the onset of the zygotic transcription by maternally derived transcription factors (TFs) such as Nanog, Pou5f1 (also known as Oct4), and Sox2 (Lee et al., 2013, Leichsenring et al., 2013), which is called zygotic genome activation (ZGA). In mammals, the massive transcriptional activation leads to the acquisition of pluripotency and the determination of cell fate such as inner cell mass (ICM) which can contribute to embryonic stem cells (ESCs) *in vitro* (Orkin et al., 2008, Yeo and Ng, 2013, Zernicka-Goetz et al., 2009).

Based on exploring the factors with transcriptional regulation in oocytes and ESCs, which can confer totipotency or pluripotency, induced pluripotent stem cells (iPSCs) were generated using TFs such as Oct4, Sox2, Klf4, c-Myc, Nanog, and Lin28a in mammals (Takahashi and Yamanaka, 2006, Yu et al., 2007). Also, higher potent two-cell (2C) embryos-like ESCs expressing murine endogenous retrovirus with leucine tRNA primer (MuERV-L) elements were characterized and induced (Macfarlan et al., 2012, Ishiuchi et al., 2015). Furthermore, extended pluripotent stem cells (EPSCs) which can contribute to both embryonic and extraembryonic lineages were established and derived by the chemical cocktail (Yang et al., 2017b, Yang et al., 2017a).

In the case of avian species, however, the partially reprogrammed iPSC-like cells have been reported using only well-known mammalian factors (Lu et al., 2012, Rossello et al., 2013, Lu et al., 2014, Choi et al., 2016, Kim et al., 2017) until recently, without discovering TFs related to very early

development in aves. In the present study, we investigated transcriptional activation and the inductive TFs at the transcriptomic level in early chicken embryos. According to the criteria of Eyal-Giladi and Kochav (EGK) (Eyal-Giladi and Kochav, 1976, Kochav et al., 1980), intrauterine development before oviposition in avian species includes ten stages, consisting of cleavage (EGK.I to EGK.VI) and the area pellucida formation period (EGK.VII to EGK.X). These two periods are clearly morphologically distinct: the former is characterized by cellularization, and the latter by axis formation and lineage segregation. During these processes, we identified TFs that are involved in regulating the two distinct periods of development. Our genome-wide expression profiling of TFs in pre-ovipositional development provides new insight into the molecular mechanisms of cellular potency and reprogramming in avian species.

## 2. Materials and methods

### *RNA-Seq data availability and pipeline for obtaining gene expression levels*

We employed our previously generated RNA-Seq data, which were deposited in the GEO database (GSE86592)(Hwang et al., 2018). The experimental use of chickens was approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1). A total of 21 raw sequencing data sets (biologically triplicated for each embryonic stage) were pre-processed using Trimmomatic (ver. 0.32) (Bolger et al., 2014). Next, clean reads were mapped onto the chicken reference genome (galGal4) obtained from Ensembl using HISAT2 (Kim et al., 2015). A Python script, HTSeq-count (Anders et al., 2014), was employed to quantify mapped reads in each Ensembl transcript.

### *Statistical test for the comparison of significantly upregulated genes between groups*

Two groups comparison were performed for consecutive stages using RNA-Seq data, and TMM normalization were used for adjusting technical variations using edgeR (Robinson et al., 2010). FDR adjustment methods were used to control for multiple comparisons, and FDR-adjusted P values of $< 0.05$ were considered statistically significant.

### *Selection of candidate DNA-dependent TFs*

We presumed that large numbers of TFs would selectively affect transcriptional transitions in early development. To investigate this, we obtained a list of TFs from the Transcription Factor DataBase (Zhang et al.,

2012). Owing to the lack of data for chickens, we also included well-known orthologous TFs from humans and mice. Consequently, 1,069 TF genes in total were compared throughout the developmental stages. The MDS method was used to assess TF expression among samples.

***Functional enrichment analysis***

For the functional investigation of the putative newly expressed genes, DAVID functional database was employed with biological databases such as GO (Huang et al., 2007), for gene set enrichment analysis of the putative newly expressed genes. We used a threshold of 5% for statistical significance.

***Exon-intron RT-PCR***

Total RNA (1 μg) was used as the template for cDNA synthesis using the SuperScript III First-Strand Synthesis System (Invitrogen). The cDNA was serially diluted five-fold and equalized quantitatively for PCR amplification. Primers for exon-intron PCR of *WNT3A* and *C8ORF22* were designed using the program Primer3 (Untergasser et al., 2012) (Table 5-9). RT-PCR was performed with an initial incubation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 59°C for 30 s, and 72°C for 30 s. The reaction was terminated after a final incubation at 72°C for 5 min.

***Correlation-based network analysis for TFs***

Based on the MDS analysis of TFs, we observed evidence for the two waves of transcriptional activation between the oocyte and zygote stages, and between EGK.III and EGK.VI. Of the candidate TFs, we filtered out

significantly detected TFs in the oocyte vs. zygote and EGK.III vs. EGK.VI comparisons. Using two significantly detected TF lists, we performed an enrichment test using DAVID based on the InterPro protein domain database (Hunter et al., 2009). Generally, a TF is included among several InterPro terms, and each of those terms includes several TFs that are similar. To identify these relationships simultaneously, we conducted a correlation-based network analysis. To convert the data, Spearman's correlation coefficients were generated a distance-matrix using a subset of TMM-normalized values. In addition, Spearman's correlation test was performed to identify significant relationships within the TF expression matrix. GO biological processes were employed for functional annotation of the detected TF genes between the oocyte and zygote stages, and between EGK.III and EGK.VI. An enrichment test (hypergeometric test) was performed for each term using DAVID (Huang et al., 2007); here, FDR-adjusted P values of $< 0.05$ were considered to indicate significantly enriched terms.

## 3. Results

### *The two waves of transcriptional activation in chicken early development*

To investigate the early embryonic TFs in chicken, we analyzed the gene expression patterns of TFs. In total, 1,069 TF genes were collected from the Animal Transcription Factor DataBase (Zhang et al., 2012). Based on the quantified TFs, stage-specific expression patterns were investigated using multidimensional scaling (MDS) analysis. A multidimensional scaling (MDS) plot showed that the TF expression profile of zygotes, clustered with EGK.I and EGK.III, is distinct from that of oocytes, suggesting that the greatest transcriptional transition between consecutive stages is observed after fertilization (Figure 5-1A). To detect newly transcribed candidates involved in transcriptional activation at the zygote stage, trimmed mean of M-value (TMM) normalized values were employed (TMM < 0 in oocyte and TMM > 0 in zygote) and 196 genes were finally identified (Table 5-1). Using this set of genes, functional enrichment analysis was performed (Figure 5-1B). The genes were significantly enriched for three GO terms: “regulation of transcription, DNA-dependent” (GO:0006355), “transcription factor complex” (GO:0005667), and “sequence-specific DNA binding” (GO:0043565).

We observed a second transcriptional transition between EGK.III and EGK.VI based on TF expression (Figure 5-1A), which is inconsistent with a previous study that identified a similar transition between late EGK.II and early EGK.III (Nagai et al., 2015). Thus, to clarify the timing of this second round of transcriptional activation, we examined the genes that were strongly activated during these stages based on the expression of primary transcripts. Exon-intron RT-PCR of two genes, *wingless-type MMTV integration site*

*family member 3A* (*WNT3A*) and *chromosome 2 open reading frame, human C8orf22* (*C8ORF22*) ($log_2$ fold-change [logFC] = 10.54 and 4.00 between EGK.III and EGK.VI, respectively), indicated that transcriptional activation occurred between EGK.IV and EGK.V, not EGK.II and EGK.III, based on actual transcripts in chicken (Figure 5-1C). These results suggest that the two waves of transcriptional changes from oocyte to zygote and from EGK.III to EGK.VI can be considered the first and second waves ($1^{st}$ and $2^{nd}$ waves) of transcriptional activation, comparable to minor and major ZGA in mammals (Xue et al., 2013). Also, transcriptional activation occurred prior to the beginning of two specific periods: zygote stage before the first cleavage, and EGK.V before area pellucida formation, respectively.

***DNA-dependent TFs activated by two waves of transcriptional activation***

Among the candidate TFs, we detected significantly activated TFs between the two transitions (oocyte to zygote and EGK.III to EGK.VI). We found that 216 and 126 TFs were activated during the first and second transitions, respectively, and that 32 TFs were activated in both transitions (false discovery rate [FDR]-adjusted $P < 0.05$ and $log_2$ fold-change [logFC] > 1; Figure 5-2A and Table 5-2~5-4). A functional enrichment analysis was performed for these TFs using the InterPro protein domain database. In total, 29 and 15 protein domain terms were significantly detected for the first and second transitions, respectively ($P < 0.05$) (Figure 5-2B and Table 5-5, 5-6). Based on TFs and protein domain terms, two waves of the transitions were compared each other (Figure 5-2A, B). As a result, the Myb-type HTH DNA-binding domain (IPR017930), Ets (IPR000418) containing *ETS1*, *ETS2*, and C-MYB TFs were exclusively up-regulated between the oocyte and zygote stages. *TCF7L2* and *LEF1*, which were clustered into the high-mobility group

HMG1/HMG2 (IPR000910), were up-regulated between EGK.III and EGK.VI only. *NANOG*, *MYC*, and *KLF4* were upregulated during both waves. Whereas, *POUV*, *SOX2*, and *LIN28A* were not detected during the first wave, but were up-regulated during the second wave. Based on these results, we conclude that specific TFs were associated with the two transition points, indicating that distinct periods of early chicken development may be differentially regulated by TFs.

***The network and expression of TFs during cleavage stages***

After identifying the TFs that were significantly activated during the two waves of transcriptional activation, we explored the function of TFs in each wave. To characterize the relationships among TFs in the first wave, TF network was constructed based on similarities in domain and co-expression patterns (Figure 5-3A). Among the core pluripotent TFs, including *NANOG, POUV*, and *SOX2*, only *NANOG* was up-regulated in the first wave and expressed throughout early development (Figure 5-3B). *POUV* and *SOX2* were not expressed during the cleavage period, and *SOX2* was expressed at lower levels than *NANOG* and *POUV* even after the second wave, indicating that ZGA may be regulated differently in chicken compared other species. Conversely, *MYC* and *KLF4* were consistently expressed during early cleavage (Figure 5-3B). TFs related to Myb-type HTH DNA-binding and the Ets domain, including *C-MYB*, *MYBL1*, *MYBL2*, *ETS1*, *ETS2*, *ETV4*, and *ETV5*, were activated after the first wave (Figure 5-3C). However, chicken-specific endogenous retrovirus *ENS-1* was not expressed during this stage. We analyzed the GO terms related to biological processes (GO-BP) during the first wave to identify the function of TFs in early development (FDR adjusted $P < 0.05$; Table 5-7). We found that chromatin remodeling (GO:0006338)

were highly enriched terms (Figure 5-3D). These results indicate that the identified TFs may be associated with cleavage stages in early chicken development.

***The functionality of TFs related to area pellucida formation period***

We established a network of similarity for TFs based on domain and expression among TFs activated after the second wave (Figure 5-4A). Homeobox TFs were clustered in this stage. Moreover, among the genes in the high-mobility group, HMG1/HMG2, Wnt-signalling-related TFs such as *TCF7L2* and *LEF1* were gradually up-regulated after EGK.VI (Figure 5-4B). As above, we analysed GO-BP among TFs up-regulated in the second wave (FDR adjusted $P < 0.05$; Table 5-8). We found that the following terms were enriched: anterior/posterior pattern specification (GO:0009952) and stem cell differentiation (GO:0048863) (Figure 5-4C). The associated TFs included many Hox genes, which are regulators of segmental identity in early development (Mallo and Alonso, 2013), and another reprogrammable *LIN28A*. Thus, the area pellucida formation period may be regulated by a group of TFs distinct from those involved in the stages of cleavage.

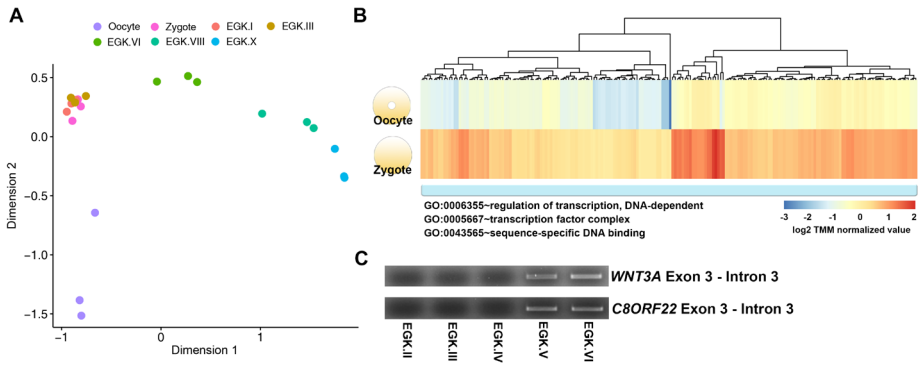


**Figure 5-1. Two waves of transcriptional activation during chicken early development.** (A) Multidimensional scaling (MDS) plot based on the expression of transcription factor genes in pre-oviposited chicken embryos. Each point represents an RNA-Seq sample, and the dotted circle represents clustered samples. First wave between oocyte and zygote; second wave between EGK.III and EGK.VI. (B) Expression and gene ontology of the newly expressed genes in zygotes (log2 TMM-normalized values < 0 in oocytes and log2 TMM-normalized values > 0 in zygotes). **(C)** Detection of gene activation via the appearance of primary transcripts based on the intronic expression of genes (*WNT3A* and *C8ORF22*) by 2$^{nd}$ ZGA with RT-PCR.

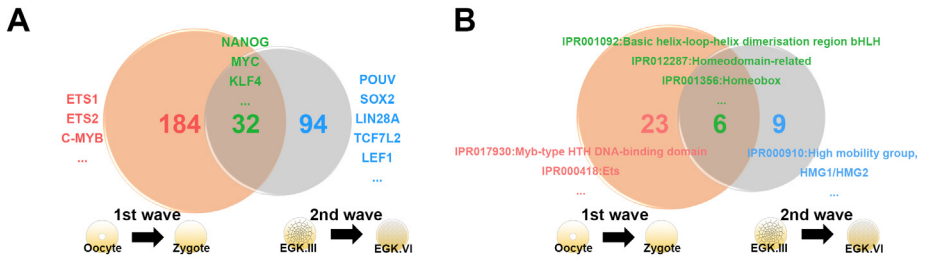


**Figure 5-2. DNA-dependent transcription factors and their domains after transcriptional activation.** (A) Venn diagram for the number of significantly detected up-regulated transcription factors (FDR adjusted $P < 0.05$ and $\log_2$ fold-change [logFC] > 1) and (B) significantly detected InterPro domains (enrichment test P-value < 0.05) between oocytes and zygotes (first wave), and between EGK.III and EGK.VI (second wave).

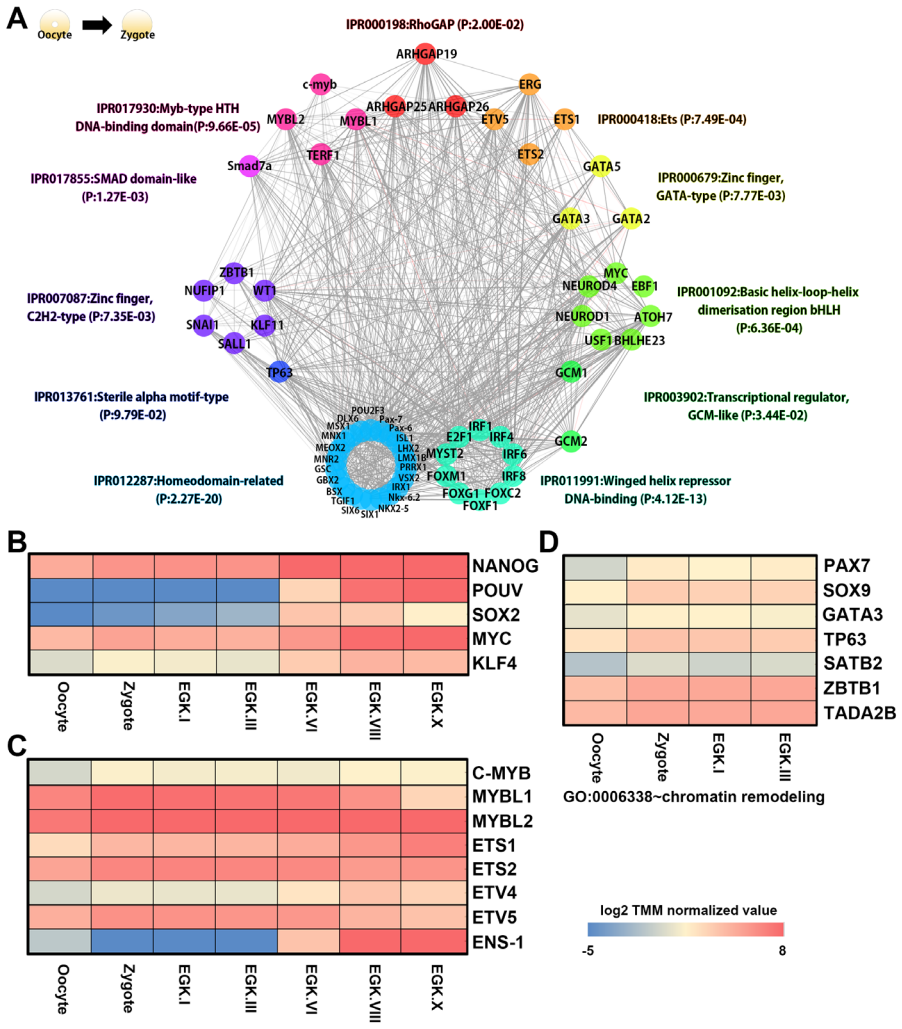


**Figure 5-3. Transcription factor network and expression from the first wave.** (A) Co-expression pattern-based transcription factor network from the first wave using the significantly up-regulated transcription factors detected between oocytes and zygotes (FDR-adjusted P < 0.05 and $\log_2$ fold-change [logFC] > 1). Each color represents significantly enriched InterPro domains. (B) Expression patterns of reprogramming-related transcription factors throughout early embryogenesis in chickens. (C) Gene expression of transcription factors related to Ets and Myb-type

HTH DNA-binding domains. (D) Gene expression of transcription factors involved in GO biological processes in chromatin remodeling.

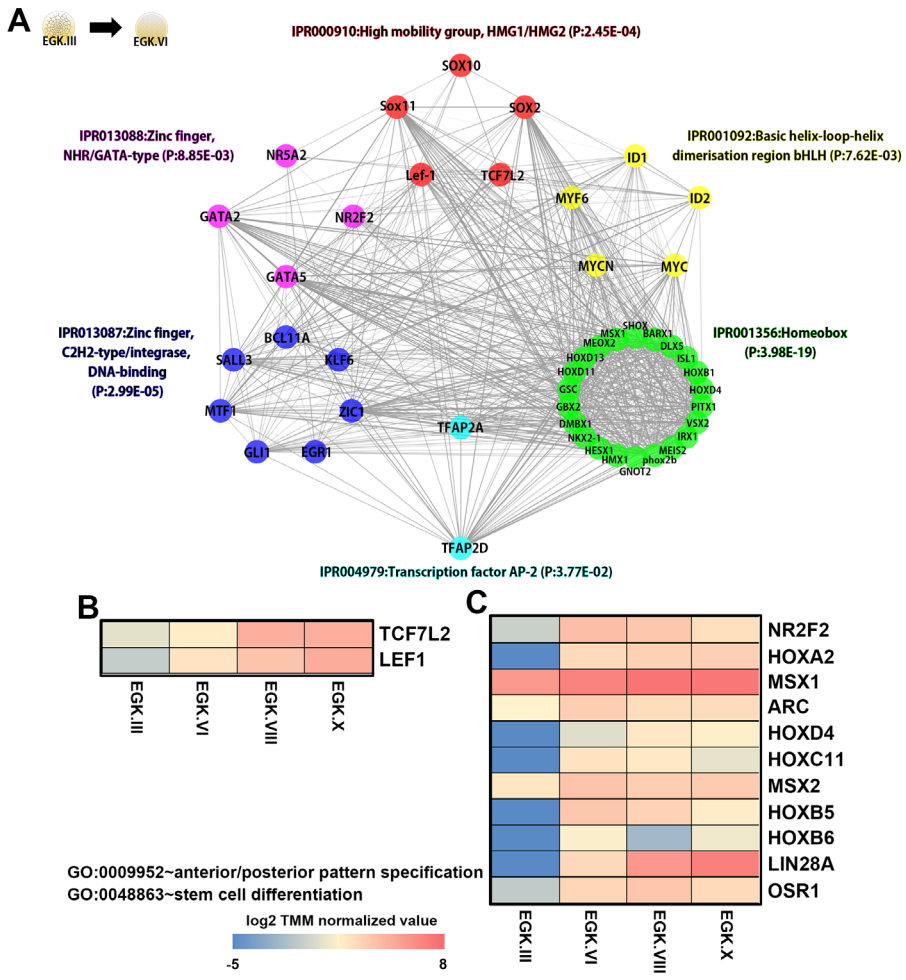


**Figure 5-4. Transcription factor network and expression from the second wave.** (A) Co-expression pattern-based transcription factor network during the 2nd wave using the significantly up-regulated transcription factors detected between EGK.III and EGK.VI (FDR-adjusted $P < 0.05$ and $\log_2$ fold-change [logFC] > 1). Each color represents significantly enriched InterPro domains. (B) Gene expression patterns of transcription factors associated with canonical Wnt signalling. (C) Heatmap representing the expression of transcription factors involved in anterior/posterior pattern specification and stem cell differentiation.

**Table 5-1. The putative newly expressed genes in zygote (log2 TMM-normalized values < 0 in oocyte and log2 TMM-normalized values > 0 in zygote)**

| Ensembl_ID | Gene_Symbol | Ensembl_ID | Gene_Symbol |
|---|---|---|---|
| ENSGALG00000000117 | TRIM27.1 | ENSGALG00000011721 | AKAP5 |
| ENSGALG00000000371 | CRHR1 | ENSGALG00000011806 | MUC13 |
| ENSGALG00000000608 | CDH1 | ENSGALG00000012123 | Pax-6 |
| ENSGALG00000000834 | RASSF5 | ENSGALG00000012433 | ADCY1 |
| ENSGALG00000000901 | CHST4 | ENSGALG00000012601 | GOLM1 |
| ENSGALG00000000936 | LMX1B | ENSGALG00000012623 | CFC1B |
| ENSGALG00000001124 | LHX2 | ENSGALG00000012768 | GCM2 |
| ENSGALG00000001214 | PARD6A | ENSGALG00000012869 | Worthington |
| ENSGALG00000001227 | PIK3R6 | ENSGALG00000012920 | Unknown |
| ENSGALG00000001352 | MORN5 | ENSGALG00000013070 | C7orf57 |
| ENSGALG00000001418 | GUCA1B | ENSGALG00000013085 | PTPRO |
| ENSGALG00000001564 | ATP2A3 | ENSGALG00000013155 | Unknown |
| ENSGALG00000001825 | FAM13C | ENSGALG00000013194 | IRX1 |
| ENSGALG00000001846 | Unknown | ENSGALG00000013223 | CLEC3A |
| ENSGALG00000001892 | CYGB | ENSGALG00000013342 | GBX2 |
| ENSGALG00000001912 | EPHA10 | ENSGALG00000013729 | NETO1 |
| ENSGALG00000002290 | Unknown | ENSGALG00000013987 | EYA4 |
| ENSGALG00000002307 | LRIT2 | ENSGALG00000014030 | CASC1 |
| ENSGALG00000002370 | SH2D4B | ENSGALG00000014820 | VTCN1 |
| ENSGALG00000002652 | FZD10 | ENSGALG00000014823 | C6orf58 |
| ENSGALG00000002840 | GRM4 | ENSGALG00000014872 | FGF10 |
| ENSGALG00000003176 | Unknown | ENSGALG00000014984 | F2RL1 |
| ENSGALG00000003193 | CRABP-I | ENSGALG00000015034 | ANKRD29 |
| ENSGALG00000003509 | C9ORF9 | ENSGALG00000015075 | Unknown |
| ENSGALG00000003690 | Unknown | ENSGALG00000015186 | KLHL14 |
| ENSGALG00000003699 | EBF1 | ENSGALG00000015410 | BVES |
| ENSGALG00000003782 | Pax-7 | ENSGALG00000015519 | ROBO2 |
| ENSGALG00000003829 | HPCAL4 | ENSGALG00000015578 | ANKRD34B |
| ENSGALG00000003876 | TIMD4 | ENSGALG00000015658 | GDAP1 |
| ENSGALG00000003895 | PRDM12 | ENSGALG00000015797 | RALYL |
| ENSGALG00000004028 | MATN4 | ENSGALG00000015845 | RIPPLY2 |
| ENSGALG00000004183 | FAM20A | ENSGALG00000015934 | Unknown |
| ENSGALG00000004257 | TSPAN15 | ENSGALG00000015941 | IL10R2 |
| ENSGALG00000004398 | CARD11 | ENSGALG00000016001 | Unknown |
| ENSGALG00000004449 | cyp3A37 | ENSGALG00000016269 | Unknown |
| ENSGALG00000004758 | KCNB1 | ENSGALG00000016326 | Unknown |
| ENSGALG00000004879 | SLC6A11 | ENSGALG00000016370 | Tpo |
| ENSGALG00000005352 | GATA5 | ENSGALG00000016577 | OTOF |
| ENSGALG00000005766 | PKD2L1 | ENSGALG00000016679 | PKHD1 |
| ENSGALG00000005928 | XAF1 | ENSGALG00000016740 | OCA2 |
| ENSGALG00000005957 | SLC6A14 | ENSGALG00000016745 | GABRB3 |
| ENSGALG00000006060 | SCN8A | ENSGALG00000016907 | KLHL1 |
| ENSGALG00000006217 | S100B | ENSGALG00000016927 | KLF5 |
| ENSGALG00000006358 | TMEM86A | ENSGALG00000017136 | GJB6 |
| ENSGALG00000006471 | SLCO2A1 | ENSGALG00000017378 | CRTAC1 |
| ENSGALG00000006480 | TCF1 | ENSGALG00000017485 | TLR1LA |

| | | | |
|---|---|---|---|
| ENSGALG00000006501 | GGT5 | ENSGALG00000018934 | ACVRL1 |
| ENSGALG00000006508 | FGF13 | ENSGALG00000019175 | Unknown |
| ENSGALG00000006515 | Protein | ENSGALG00000019384 | MAPK12 |
| ENSGALG00000006565 | Unknown | ENSGALG00000020143 | GABRG1 |
| ENSGALG00000006665 | GALNT6 | ENSGALG00000020332 | Unknown |
| ENSGALG00000006673 | POU2F3 | ENSGALG00000020609 | Unknown |
| ENSGALG00000006795 | ABCG4 | ENSGALG00000020625 | Unknown |
| ENSGALG00000006981 | MCTP2 | ENSGALG00000020905 | KSR2 |
| ENSGALG00000007028 | Unknown | ENSGALG00000021686 | Unknown |
| ENSGALG00000007080 | Unknown | ENSGALG00000022237 | Unknown |
| ENSGALG00000007171 | CIITA | ENSGALG00000022644 | CCDC36 |
| ENSGALG00000007211 | CDH22 | ENSGALG00000022940 | FOXC2 |
| ENSGALG00000007288 | SNTN | ENSGALG00000022976 | OFCC1 |
| ENSGALG00000007448 | SPRY3 | ENSGALG00000023195 | DLX6 |
| ENSGALG00000007623 | CACNA1G | ENSGALG00000023271 | Unknown |
| ENSGALG00000007631 | MEGF11 | ENSGALG00000023441 | RTN4RL2 |
| ENSGALG00000007815 | ANKK1 | ENSGALG00000023444 | Unknown |
| ENSGALG00000007817 | EHF | ENSGALG00000023587 | Unknown |
| ENSGALG00000007892 | SPAG6 | ENSGALG00000023961 | GDF8 |
| ENSGALG00000008085 | NEFH | ENSGALG00000024024 | TGFA |
| ENSGALG00000008135 | SATB2 | ENSGALG00000024243 | Unknown |
| ENSGALG00000008177 | NOS1 | ENSGALG00000024396 | NEUROD4 |
| ENSGALG00000008262 | RASGRF1 | ENSGALG00000025721 | Unknown |
| ENSGALG00000008462 | CDK3 | ENSGALG00000025723 | NKX1-1 |
| ENSGALG00000008735 | BFSP1 | ENSGALG00000026130 | Unknown |
| ENSGALG00000008908 | NEUROD1 | ENSGALG00000026264 | MAFA |
| ENSGALG00000008992 | ETNPPL | ENSGALG00000026360 | SOWAHC |
| ENSGALG00000009268 | FGG | ENSGALG00000026391 | Unknown |
| ENSGALG00000009487 | Unknown | ENSGALG00000026510 | Unknown |
| ENSGALG00000009520 | MARCH1 | ENSGALG00000026651 | SCARA3 |
| ENSGALG00000009589 | GAD67 | ENSGALG00000026698 | Unknown |
| ENSGALG00000009616 | SPATA17 | ENSGALG00000026863 | Unknown |
| ENSGALG00000009628 | Unknown | ENSGALG00000026881 | PITX2 |
| ENSGALG00000009719 | PCDH10 | ENSGALG00000027036 | PKIB |
| ENSGALG00000009791 | PROX1 | ENSGALG00000027075 | RAB25 |
| ENSGALG00000009853 | FOXG1 | ENSGALG00000027121 | Unknown |
| ENSGALG00000009865 | RNF150 | ENSGALG00000027235 | Unknown |
| ENSGALG00000010148 | PTPRR | ENSGALG00000027255 | NRXN3 |
| ENSGALG00000010249 | Unknown | ENSGALG00000027351 | LRRC71 |
| ENSGALG00000010323 | BATF | ENSGALG00000027419 | Unknown |
| ENSGALG00000010342 | OTOGL | ENSGALG00000027511 | SP5 |
| ENSGALG00000010486 | JAKMIP3 | ENSGALG00000027525 | Unknown |
| ENSGALG00000010487 | Unknown | ENSGALG00000027630 | NPR2 |
| ENSGALG00000010599 | SULT6B1 | ENSGALG00000027765 | Unknown |
| ENSGALG00000010662 | Unknown | ENSGALG00000027885 | C4orf19 |
| ENSGALG00000010974 | GSC | ENSGALG00000028283 | Unknown |
| ENSGALG00000011009 | SCN2A | ENSGALG00000028388 | Unknown |
| ENSGALG00000011157 | TRIM67 | ENSGALG00000028560 | OC3 |
| ENSGALG00000011262 | KCNH8 | ENSGALG00000028589 | DMTN |
| ENSGALG00000011270 | KERA | ENSGALG00000028826 | ABHD8 |
| ENSGALG00000011344 | TTLL2 | ENSGALG00000029012 | Unknown |
| ENSGALG00000011544 | BCL2L14 | ENSGALG00000029023 | DMRTB1 |

**Table 5-2. List of significantly upregulated transcription factors during both 1st ZGA and 2nd ZGA (FDR adjusted P-value < 0.05 and log2 fold-change (logFC) > 1)**

| Ensembl_ID | Gene_Symbol | Description |
|---|---|---|
| ENSGALG00000001188 | ETV4 | ets variant 4 |
| ENSGALG00000005352 | GATA5 | transcription factor GATA-5 |
| ENSGALG00000005909 | GATA2 | GATA-binding factor 2 |
| ENSGALG00000007690 | TFAP2C | transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma) |
| ENSGALG00000008253 | TBX5 | T-box transcription factor TBX5 |
| ENSGALG00000008702 | OVOL2 | ovo-like zinc finger 2 |
| ENSGALG00000010101 | Unknown | Unknown |
| ENSGALG00000010213 | VSX2 | Visual system homeobox 2 |
| ENSGALG00000010794 | MEOX2 | homeobox protein MOX-2 |
| ENSGALG00000010974 | GSC | homeobox protein goosecoid |
| ENSGALG00000012775 | TFAP2A | Gallus gallus transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha) (TFAP2A), mRNA. |
| ENSGALG00000013194 | IRX1 | iroquois-class homeodomain protein IRX-1 |
| ENSGALG00000013342 | GBX2 | Gallus gallus gastrulation brain homeobox 2 (GBX2), mRNA. |
| ENSGALG00000014208 | GRHL2 | grainyhead-like transcription factor 2 |
| ENSGALG00000014558 | ZNF774 | zinc finger protein 774 |
| ENSGALG00000014884 | ISL1 | Insulin gene enhancer protein ISL-1 |
| ENSGALG00000015013 | MSX1 | Homeobox protein GHOX-7 |
| ENSGALG00000016308 | MYC | Myc proto-oncogene protein |
| ENSGALG00000016397 | Sox11 | transcription factor SOX-11 |
| ENSGALG00000016927 | KLF5 | Kruppel-like factor 5 (intestinal) |
| ENSGALG00000019025 | UNCX | UNC homeobox |
| ENSGALG00000019470 | MSC | musculin |

| | | |
|---|---|---|
| ENSGALG00000022809 | PRDM14 | PR domain containing 14 |
| ENSGALG00000023430 | EN2 | Homeobox protein engrailed-2 |
| ENSGALG00000025723 | NKX1-1 | NK1 homeobox 1 |
| ENSGALG00000026180 | MSX2 | Homeobox protein MSX-2 |
| ENSGALG00000026486 | ZNF322 | zinc finger protein 322 |
| ENSGALG00000026928 | KLF4 | Kruppel-like factor 4 (gut) |
| ENSGALG00000027247 | EOMES | eomesodermin |
| ENSGALG00000027498 | FOXB1 | forkhead box B1 |
| ENSGALG00000027720 | OTX2 | homeobox protein OTX2 |
| ENSGALG00000027772 | NANOG | nanog homeobox |

**Table 5-3. List of significantly upregulated transcription factors during 1st ZGA only (FDR adjusted P-value < 0.05 and log2 fold-change (logFC) > 1)**

| Ensembl_ID | Gene_Symbol | Description |
|---|---|---|
| ENSGALG00000000043 | USF1 | Gallus gallus upstream transcription factor 1 (USF1), mRNA. |
| ENSGALG00000000132 | ARHGAP25 | rho GTPase-activating protein 25 |
| ENSGALG00000000164 | MYBPH | Myosin-binding protein H |
| ENSGALG00000000233 | SP2 | Sp2 transcription factor |
| ENSGALG00000000373 | Unknown | Uncharacterized protein |
| ENSGALG00000000725 | RFX7 | regulatory factor X, 7 |
| ENSGALG00000000838 | PRKRIR | protein-kinase, interferon-inducible double stranded RNA dependent inhibitor, repressor of (P58 repressor) |
| ENSGALG00000000936 | LMX1B | LIM/homeobox protein LMX-1.2 |
| ENSGALG00000001006 | TP73 | tumor protein p73 |
| ENSGALG00000001014 | ARPC5L | actin related protein 2/3 complex, subunit 5-like |
| ENSGALG00000001058 | ARHGEF16 | Rho guanine nucleotide exchange factor (GEF) 16 |
| ENSGALG00000001124 | LHX2 | LIM homeobox 2 |
| ENSGALG00000001143 | ETS1 | protein C-ets-1 |
| ENSGALG00000001311 | ARHGAP44 | Rho GTPase activating protein 44 |
| ENSGALG00000001376 | ARID3B | AT-rich interactive domain-containing protein 3B |
| ENSGALG00000001403 | HIC2 | hypermethylated in cancer 2 protein |
| ENSGALG00000001405 | IRF6 | interferon regulatory factor 6 |
| ENSGALG00000001735 | LHX4 | LIM homeobox 4 |
| ENSGALG00000001836 | PIAS4 | protein inhibitor of activated STAT, 4 |
| ENSGALG00000001910 | SNAPC4 | small nuclear RNA activating complex, polypeptide 4, 190kDa |
| ENSGALG00000002528 | DNAJC18 | dnaJ homolog subfamily C member 18 |
| ENSGALG00000002874 | ZBTB38 | zinc finger and BTB domain containing 38 |
| ENSGALG00000002886 | NKX2-5 | homeobox protein Nkx-2.5 |

| | | |
|---|---|---|
| ENSGALG00000003038 | Unknown | Uncharacterized protein |
| ENSGALG00000003045 | E2F1 | transcription factor E2F1 |
| ENSGALG00000003217 | LITAF | lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog |
| ENSGALG00000003324 | PRRX1 | paired mesoderm homeobox protein 1 |
| ENSGALG00000003387 | DNAJC21 | DnaJ (Hsp40) homolog, subfamily C, member 21 |
| ENSGALG00000003406 | RXRG | retinoic acid receptor RXR-gamma |
| ENSGALG00000003424 | LMX1A | LIM homeobox transcription factor 1-alpha |
| ENSGALG00000003503 | MYBL2 | myb-related protein B |
| ENSGALG00000003699 | EBF1 | transcription factor COE1 |
| ENSGALG00000003737 | TOX3 | TOX high mobility group box family member 3 isoform 1 |
| ENSGALG00000003739 | SALL1 | sal-like protein 1 |
| ENSGALG00000003782 | Pax-7 | paired box protein Pax-7 |
| ENSGALG00000003895 | PRDM12 | PR domain containing 12 |
| ENSGALG00000003931 | ATOH7 | Protein atonal homolog 7 |
| ENSGALG00000003939 | KLF2 | Kruppel-like factor 2 |
| ENSGALG00000003952 | RBPJL | recombination signal binding protein for immunoglobulin kappa J region-like |
| ENSGALG00000004023 | E2F8 | E2F transcription factor 8 |
| ENSGALG00000004386 | Sox9 | Transcription factor SOX-9 |
| ENSGALG00000004477 | NCOR1 | nuclear receptor corepressor 1 |
| ENSGALG00000004532 | ZNF507 | zinc finger protein 507 |
| ENSGALG00000004598 | CUX2 | cut-like homeobox 2 |
| ENSGALG00000004663 | ARFGEF2 | ADP-ribosylation factor guanine nucleotide-exchange factor 2 (brefeldin A-inhibited) |
| ENSGALG00000005441 | NFIB | nuclear factor 1 B-type |
| ENSGALG00000005635 | TCFL5 | transcription factor-like 5 (basic helix-loop-helix) |
| ENSGALG00000005683 | ARHGAP29 | rho GTPase-activating protein 29 |
| ENSGALG00000005718 | BHLHE23 | class E basic helix-loop-helix protein 23 |
| ENSGALG000 | Unknown | Unknown |

| | | |
|---|---|---|
| 00005728 | | |
| ENSGALG00000005757 | IRF8 | Interferon regulatory factor 8 |
| ENSGALG00000005768 | ETV5 | ets variant 5 |
| ENSGALG00000005932 | MYT1 | myelin transcription factor 1 |
| ENSGALG00000006062 | ZBTB46 | zinc finger and BTB domain containing 46 |
| ENSGALG00000006099 | ZFPM1 | zinc finger protein, FOG family member 1 |
| ENSGALG00000006137 | ARHGAP22 | Rho GTPase activating protein 22 |
| ENSGALG00000006438 | Unknown | Unknown |
| ENSGALG00000006439 | ARHGEF6 | Rho guanine nucleotide exchange factor 6 |
| ENSGALG00000006480 | TCF1 | transcription factor 7 |
| ENSGALG00000006497 | ZIC3 | Zic family member 3 |
| ENSGALG00000006515 | Protein | Brain-specific homeobox protein homolog |
| ENSGALG00000006663 | ARHGEF12 | Rho guanine nucleotide exchange factor (GEF) 12 |
| ENSGALG00000006673 | POU2F3 | POU class 2 homeobox 3 |
| ENSGALG00000006726 | GATA3 | GATA-binding factor 3 |
| ENSGALG00000006785 | IRF1 | Interferon regulatory factor 1 |
| ENSGALG00000007296 | MAFG | Transcription factor MafG |
| ENSGALG00000007324 | TP63 | tumor protein 63 |
| ENSGALG00000007380 | ARHGAP26 | Rho GTPase-activating protein 26 |
| ENSGALG00000007570 | Unknown | Uncharacterized protein |
| ENSGALG00000007609 | ARHGAP19 | rho GTPase-activating protein 19 |
| ENSGALG00000007817 | EHF | ETS homologous factor |
| ENSGALG00000007834 | SALL4 | sal-like protein 4 |
| ENSGALG00000007973 | F | nuclear respiratory factor 1 |
| ENSGALG00000008014 | CEBPB | CCAAT/enhancer-binding protein beta |
| ENSGALG00000008018 | SNAI1 | snail homolog 1 |
| ENSGALG00000008049 | ARL3 | ADP-ribosylation factor-like 3 |

| | | |
|---|---|---|
| ENSGALG00000008135 | SATB2 | DNA-binding protein SATB2 |
| ENSGALG00000008303 | LHX5 | LIM homeobox 5 |
| ENSGALG00000008393 | CREB3L1 | cAMP responsive element binding protein 3-like 1 |
| ENSGALG00000008501 | KLF7 | Kruppel-like factor 7 (ubiquitous) |
| ENSGALG00000008653 | MGA | MGA, MAX dimerization protein |
| ENSGALG00000008657 | TUBGCP6 | tubulin, gamma complex associated protein 6 |
| ENSGALG00000008774 | MEIS1 | Meis homeobox 1 |
| ENSGALG00000008908 | NEUROD1 | neurogenic differentiation factor 1 |
| ENSGALG00000009012 | ZNF385B | zinc finger protein 385B |
| ENSGALG00000009081 | ZNF106 | zinc finger protein 106 |
| ENSGALG00000009168 | AFF2 | AF4/FMR2 family, member 2 |
| ENSGALG00000009302 | EMX2 | empty spiracles homeobox 2 |
| ENSGALG00000009339 | TP53BP2 | tumor protein p53 binding protein 2 |
| ENSGALG00000009437 | MECOM | MDS1 and EVI1 complex locus |
| ENSGALG00000009791 | PROX1 | prospero homeobox protein 1 |
| ENSGALG00000009853 | FOXG1 | forkhead box protein G1 |
| ENSGALG00000009944 | MYST2 | histone acetyltransferase MYST2 |
| ENSGALG00000010035 | NR3C2 | Mineralocorticoid receptor |
| ENSGALG00000010265 | E2F7 | E2F transcription factor 7 |
| ENSGALG00000010320 | PAWR | PRKC, apoptosis, WT1, regulator |
| ENSGALG00000010323 | BATF | basic leucine zipper transcription factor, ATF-like |
| ENSGALG00000010500 | Nkx-6.2 | homeobox protein Nkx-6.2 |
| ENSGALG00000010905 | SP4 | Sp4 transcription factor |
| ENSGALG00000011349 | Unknown | Uncharacterized protein |
| ENSGALG00000011390 | REST | RE1-silencing transcription factor |
| ENSGALG00000011393 | ZNF142 | zinc finger protein 142 |
| ENSGALG000 | ARHGAP35 | Rho GTPase activating protein 35 |

| | | |
|---|---|---|
| 00011693 | | |
| ENSGALG00000011718 | ZBTB1 | zinc finger and BTB domain containing 1 |
| ENSGALG00000011722 | STRA8 | stimulated by retinoic acid 8 |
| ENSGALG00000011738 | ARHGDIB | rho GDP-dissociation inhibitor 2 |
| ENSGALG00000012046 | ARPP21 | cAMP-regulated phosphoprotein, 21kDa |
| ENSGALG00000012115 | WT1 | Wilms tumor 1 |
| ENSGALG00000012123 | Pax-6 | Paired box protein Pax-6 |
| ENSGALG00000012199 | WDHD1 | WD repeat and HMG-box DNA binding protein 1 |
| ENSGALG00000012556 | HMGXB4 | HMG box domain containing 4 |
| ENSGALG00000012768 | GCM2 | chorion-specific transcription factor GCMb |
| ENSGALG00000012830 | IRF-4 | interferon regulatory factor 4 |
| ENSGALG00000012958 | ZSCAN2 | zinc finger and SCAN domain containing 2 |
| ENSGALG00000013056 | TUBA8 | tubulin, alpha 8 |
| ENSGALG00000013420 | FOXM1 | forkhead box protein M1 |
| ENSGALG00000013683 | ARID1B | AT rich interactive domain 1B (SWI1-like) |
| ENSGALG00000013754 | PLAGL1 | pleiomorphic adenoma gene-like 1 |
| ENSGALG00000013956 | c-myb | transcriptional activator Myb |
| ENSGALG00000014795 | TGIF1 | Homeobox protein AKR |
| ENSGALG00000014808 | KIAA2018 | uncharacterized protein LOC418320 |
| ENSGALG00000014825 | Unknown | Uncharacterized protein |
| ENSGALG00000014918 | RFX6 | regulatory factor X, 6 |
| ENSGALG00000015163 | RNF138 | ring finger protein 138, E3 ubiquitin protein ligase |
| ENSGALG00000015289 | ARMC2 | armadillo repeat containing 2 |
| ENSGALG00000015522 | MYBL1 | myb-related protein A |
| ENSGALG00000015547 | TADA2B | transcriptional adaptor 2B |
| ENSGALG00000015605 | BACH2 | BTB and CNC homology 1, basic leucine zipper transcription factor 2 |
| ENSGALG00000015633 | TERF1 | telomeric repeat-binding factor 1 |

| | | |
|---|---|---|
| ENSGALG00000015687 | ZFYVE28 | zinc finger, FYVE domain containing 28 |
| ENSGALG00000015705 | WHSC1 | Wolf-Hirschhorn syndrome candidate 1 |
| ENSGALG00000015834 | TBX18 | Gallus gallus T-box 18 (TBX18), mRNA. |
| ENSGALG00000016058 | ERG | Transcriptional regulator Erg |
| ENSGALG00000016059 | ETS2 | Protein C-ets-2 |
| ENSGALG00000016316 | GCM1 | chorion-specific transcription factor GCMa |
| ENSGALG00000016439 | GRHL1 | grainyhead-like transcription factor 1 |
| ENSGALG00000016440 | KLF11 | Krueppel-like factor 11 |
| ENSGALG00000016506 | ZNF512 | zinc finger protein 512 |
| ENSGALG00000016915 | MYCBP2 | MYC binding protein 2, E3 ubiquitin protein ligase |
| ENSGALG00000016972 | NUFIP1 | nuclear fragile X mental retardation-interacting protein 1 |
| ENSGALG00000017276 | ARHGEF5 | Rho guanine nucleotide exchange factor (GEF) 5 |
| ENSGALG00000017389 | SIX4 | SIX homeobox 4 |
| ENSGALG00000018503 | NKX6-3 | NK6 homeobox 3 |
| ENSGALG00000018639 | Smad7a | TGF-beta signal pathway antagonist Smad7 |
| ENSGALG00000019041 | NOBOX | NOBOX oogenesis homeobox |
| ENSGALG00000019647 | Unknown | Unknown |
| ENSGALG00000019711 | Unknown | Unknown |
| ENSGALG00000020359 | SIX6 | Homeobox protein SIX6 |
| ENSGALG00000021593 | HEYL | hes-related family bHLH transcription factor with YRPW motif-like |
| ENSGALG00000022935 | FOXI3 | forkhead box I3 |
| ENSGALG00000022940 | FOXC2 | Gallus gallus forkhead box C2 (MFH-1, mesenchyme forkhead 1) (FOXC2), mRNA. |
| ENSGALG00000022994 | SIX1 | homeobox protein SIX1 |
| ENSGALG00000023025 | ZBTB42 | zinc finger and BTB domain containing 42 |
| ENSGALG00000023195 | DLX6 | distal-less homeobox 6 |
| ENSGALG00000023406 | NKX1-2 | NK1 homeobox 2 |
| ENSGALG000 | TGIF2 | TGFB-induced factor homeobox 2 |

| | | |
|---|---|---|
| 00024081 | | |
| ENSGALG00000024396 | NEUROD4 | neurogenic differentiation factor 4 |
| ENSGALG00000025759 | TFDP2 | transcription factor Dp-2 |
| ENSGALG00000025795 | ARHGAP23 | Rho GTPase activating protein 23 |
| ENSGALG00000025911 | ARHGDIG | Rho GDP dissociation inhibitor (GDI) gamma |
| ENSGALG00000026243 | POU4F3 | Brain-specific homeobox/POU domain protein 3 |
| ENSGALG00000026264 | MAFA | transcription factor MafA |
| ENSGALG00000026276 | TCF15 | transcription factor 15 |
| ENSGALG00000026322 | ARMT1 | acidic residue methyltransferase 1 |
| ENSGALG00000026582 | SOX18 | transcription factor SOX-18 |
| ENSGALG00000026653 | ARL9 | ADP-ribosylation factor-like 9 |
| ENSGALG00000026710 | DNAJC28 | DnaJ (Hsp40) homolog, subfamily C, member 28 |
| ENSGALG00000026761 | LIN28B | Protein lin-28 homolog B |
| ENSGALG00000026792 | E2F2 | E2F transcription factor 2 |
| ENSGALG00000026881 | PITX2 | pituitary homeobox 2 |
| ENSGALG00000027335 | ARID5A | AT rich interactive domain 5A (MRF1-like) |
| ENSGALG00000027381 | OLIG3 | Gallus gallus oligodendrocyte transcription factor 3 (OLIG3), mRNA. |
| ENSGALG00000027511 | SP5 | Sp5 transcription factor |
| ENSGALG00000027677 | DLX3 | distal-less homeobox 3 |
| ENSGALG00000027684 | TUBB6 | Tubulin beta-5 chain |
| ENSGALG00000027993 | MXD3 | MAX dimerization protein 3 |
| ENSGALG00000028081 | ONECUT3 | one cut homeobox 3 |
| ENSGALG00000028360 | MGAT4C | MGAT4 family, member C |
| ENSGALG00000028576 | PAX1 | paired box 1 |
| ENSGALG00000028601 | ZBED1 | zinc finger, BED-type containing 1 |
| ENSGALG00000029023 | DMRTB1 | doublesex- and mab-3-related transcription factor B1 |
| ENSGALG00000029036 | ZNF687 | zinc finger protein 687 |

| | | |
|---|---|---|
| ENSGALG00000029154 | GATAD2B | GATA zinc finger domain containing 2B |
| ENSGALG00000029178 | ZNF628 | zinc finger protein 628 |

**Table 5-4. List of significantly upregulated transcription factors during 2nd ZGA only (FDR adjusted P-value < 0.05 and log2 fold-change (logFC) > 1)**

| Ensembl_ID | Gene_Symbol | Description |
|---|---|---|
| ENSGALG00000000177 | KLF3 | Kruppel-like factor 3 (basic) |
| ENSGALG00000000388 | LIN28A | protein lin-28 homolog A |
| ENSGALG00000000406 | Unknown | Uncharacterized protein |
| ENSGALG00000000583 | SOX13 | SRY (sex determining region Y)-box 13 |
| ENSGALG00000000616 | ELF3 | E74-like factor 3 (ets domain transcription factor, epithelial-specific ) |
| ENSGALG00000000934 | ZBTB43 | zinc finger and BTB domain containing 43 |
| ENSGALG00000001101 | MBD3 | methyl-CpG-binding domain protein 3 |
| ENSGALG00000001161 | FLI1 | Friend leukemia integration 1 transcription factor |
| ENSGALG00000001276 | HOXB9 | homeobox B9 |
| ENSGALG00000001518 | ZNF750 | zinc finger protein 750 |
| ENSGALG00000001639 | MTF1 | metal regulatory transcription factor 1 |
| ENSGALG00000002182 | NR5A2 | nuclear receptor subfamily 5 group A member 2 |
| ENSGALG00000003589 | VTN | vitronectin precursor |
| ENSGALG00000004994 | ZNF703 | zinc finger protein 703 |
| ENSGALG00000005131 | BARX1 | homeobox protein BarH-like 1b |
| ENSGALG00000005506 | HESX1 | homeobox protein ANF-1 |
| ENSGALG00000005613 | Unknown | Unknown |
| ENSGALG00000005689 | PAX2 | paired box 2 |
| ENSGALG00000006210 | Id1 | DNA-binding protein inhibitor ID-1 |
| ENSGALG00000006241 | TUBA1C | tubulin, alpha 1c |
| ENSGALG00000006365 | PITX1 | Pituitary homeobox 1 |
| ENSGALG00000006448 | IRF10 | interferon regulatory factor 10 |
| ENSGALG00000006811 | ZIC1 | zinc finger protein ZIC 1 |

| | | |
|---|---|---|
| ENSGALG00000006933 | C16orf52 | chromosome 16 open reading frame 52 |
| ENSGALG00000007000 | NR2F2 | COUP transcription factor 2 |
| ENSGALG00000007056 | KLF6 | Gallus gallus Kruppel-like factor 6 (KLF6), mRNA. |
| ENSGALG00000007184 | FEZF2 | Uncharacterized protein |
| ENSGALG00000007608 | ARHGEF9 | Cdc42 guanine nucleotide exchange factor (GEF) 9 |
| ENSGALG00000007669 | EGR1 | Early growth response protein 1 |
| ENSGALG00000007866 | BCL11A | B-cell lymphoma/leukemia 11A |
| ENSGALG00000008148 | MYBPC3 | myosin binding protein C, cardiac |
| ENSGALG00000008848 | SOX2 | Transcription factor SOX-2 |
| ENSGALG00000008883 | TCF7L2 | transcription factor 7-like 2 |
| ENSGALG00000009129 | DLX5 | Homeobox protein DLX-5 |
| ENSGALG00000009273 | HOXD11 | Homeobox protein Hox-D11 |
| ENSGALG00000009277 | HOXD13 | Gallus gallus homeobox D13 (HOXD13), mRNA. |
| ENSGALG00000009799 | MEIS2 | Meis homeobox 2 |
| ENSGALG00000010160 | DMRT1 | doublesex- and mab-3-related transcription factor 1 |
| ENSGALG00000010161 | DMRT3 | doublesex and mab-3 related transcription factor 3 |
| ENSGALG00000010436 | DMBX1 | diencephalon/mesencephalon homeobox 1 |
| ENSGALG00000010461 | EBF3 | early B-cell factor 3 |
| ENSGALG00000010529 | Lef-1 | lymphoid enhancer-binding factor 1 |
| ENSGALG00000010577 | ARHGEF38 | Rho guanine nucleotide exchange factor (GEF) 38 |
| ENSGALG00000010588 | Unknown | CCR4-NOT transcription complex, subunit 2 |
| ENSGALG00000010880 | Unknown | Uncharacterized protein |
| ENSGALG00000010936 | MYF6 | Myogenic factor 6 |
| ENSGALG00000011053 | HOXA2 | Homeobox protein Hox-A2 |
| ENSGALG00000011137 | CREB5 | cAMP responsive element binding protein 5 |
| ENSGALG00000011657 | EAF2 | ELL-associated factor 2 |
| ENSGALG000000 | SOX10 | Transcription factor SOX-10 |

| | | |
|---|---|---|
| 12290 | | |
| ENSGALG00000012557 | ISX | intestine-specific homeobox |
| ENSGALG00000012629 | ZNF367 | zinc finger protein 367 |
| ENSGALG00000012657 | SALL3 | sal-like protein 3 |
| ENSGALG00000013454 | HOXB1 | homeobox protein Hox-B1 |
| ENSGALG00000013568 | NR4A3 | nuclear receptor subfamily 4, group A, member 3 |
| ENSGALG00000014250 | phox2b | Uncharacterized protein |
| ENSGALG00000014639 | Unknown | Unknown |
| ENSGALG00000014697 | MADH2 | mothers against decapentaplegic homolog 2 |
| ENSGALG00000015112 | ZNF521 | zinc finger protein 521 |
| ENSGALG00000015388 | Blimp-1 | Uncharacterized protein |
| ENSGALG00000015600 | HMX1 | Homeobox protein HMX1 |
| ENSGALG00000015670 | HNF4G | hepatocyte nuclear factor 4, gamma |
| ENSGALG00000015926 | RUNX1T1 | protein CBFA2T1 |
| ENSGALG00000015987 | TP53INP1 | tumor protein p53-inducible nuclear protein 1 |
| ENSGALG00000016154 | ARC | Activity-regulated cytoskeleton-associated protein |
| ENSGALG00000016403 | ID2 | DNA-binding protein inhibitor ID-2 |
| ENSGALG00000016462 | MYCN | N-myc proto-oncogene protein |
| ENSGALG00000016473 | OSR1 | protein odd-skipped-related 1 |
| ENSGALG00000016698 | SHOX | short stature homeobox protein |
| ENSGALG00000017503 | Unknown | Uncharacterized protein |
| ENSGALG00000018733 | Unknown | Unknown |
| ENSGALG00000018783 | HOXC10 | homeobox C10 |
| ENSGALG00000019842 | TFAP2D | transcription factor AP-2 delta (activating enhancer binding protein 2 delta) |
| ENSGALG00000020423 | NKX2-1 | NK2 homeobox 1 |
| ENSGALG00000020522 | MYRF | myelin regulatory factor |
| ENSGALG00000020627 | SHOX2 | short stature homeobox 2 |

| | | |
|---|---|---|
| ENSGALG00000022295 | Unknown | Uncharacterized protein |
| ENSGALG00000023419 | HOXD4 | homeobox protein Hox-D4 |
| ENSGALG00000025887 | ATF3 | activating transcription factor 3 |
| ENSGALG00000026100 | HOXC11 | homeobox C11 |
| ENSGALG00000026393 | HOXC12 | homeobox C12 |
| ENSGALG00000026574 | POUV | POU domain, class 5, transcription factor 1 |
| ENSGALG00000026744 | OTP | orthopedia homeobox |
| ENSGALG00000026758 | SOX2OT_exon 4 | SOX2 overlapping transcript exon 4 |
| ENSGALG00000026787 | HAND1 | heart- and neural crest derivatives-expressed protein 1 |
| ENSGALG00000027242 | TTF1 | transcription termination factor, RNA polymerase I |
| ENSGALG00000027365 | HIVEP3 | human immunodeficiency virus type I enhancer binding protein 3 |
| ENSGALG00000027554 | HLX | H2.0-like homeobox |
| ENSGALG00000027764 | CDX2 | caudal type homeobox 2 |
| ENSGALG00000027903 | HOXB5 | Homeobox protein Hox-B5 |
| ENSGALG00000028276 | NEUROD2 | neuronal differentiation 2 |
| ENSGALG00000028483 | NKX2-8 | NK2 homeobox 8 |
| ENSGALG00000028726 | HOXB6 | homeobox B6 |
| ENSGALG00000028950 | Six3os1_5 | Six3os1 conserved region 5 |

**Table 5-5. Significantly enriched protein domain terms in upregulated transcription factors during 1st ZGA (enrichment test P-value < 0.05)**

| Term | P-Value |
|---|---|
| IPR012287:Homeodomain-related | 2.27E-20 |
| IPR001356:Homeobox | 5.42E-16 |
| IPR017970:Homeobox, conserved site | 2.74E-15 |
| IPR011991:Winged helix repressor DNA-binding | 4.12E-13 |
| IPR017930:Myb-type HTH DNA-binding domain | 9.66E-05 |
| IPR019817:Interferon regulatory factor, conserved site | 1.67E-04 |
| IPR001346:Interferon regulatory factor | 2.64E-04 |
| IPR001092:Basic helix-loop-helix dimerisation region bHLH | 6.36E-04 |
| IPR014778:Myb, DNA-binding | 7.49E-04 |
| IPR000418:Ets | 7.49E-04 |
| IPR015395:C-myb, C-terminal | 8.80E-04 |
| IPR001005:SANT, DNA-binding | 9.86E-04 |
| IPR011598:Helix-loop-helix DNA-binding | 1.20E-03 |
| IPR017855:SMAD domain-like | 1.27E-03 |
| IPR015495:Myb transcription factor | 1.74E-03 |
| IPR018122:Transcription factor, fork head, conserved site | 1.96E-03 |
| IPR001766:Transcription factor, fork head | 1.96E-03 |
| IPR015880:Zinc finger, C2H2-like | 2.00E-03 |
| IPR003118:Sterile alpha motif/pointed | 4.26E-03 |
| IPR019471:Interferon regulatory factor-3 | 4.26E-03 |
| IPR007087:Zinc finger, C2H2-type | 7.35E-03 |
| IPR000679:Zinc finger, GATA-type | 7.77E-03 |
| IPR013761:Sterile alpha motif-type | 9.79E-03 |
| IPR000198:RhoGAP | 2.05E-02 |
| IPR016374:Transcription factor, GATA-1/2/3 | 3.44E-02 |
| IPR016311:Transforming factor C-ets | 3.44E-02 |
| IPR012642:Transcriptional regulator, Wos2-domain | 3.44E-02 |
| IPR016637:Transcription factor, basic helix-loop-helix, NeuroD | 3.44E-02 |
| IPR003902:Transcriptional regulator, GCM-like | 3.44E-02 |

**Table 5-6. Significantly enriched protein domain terms in upregulated transcription factors during 2nd ZGA (enrichment test P-value < 0.05)**

| Term | P-Value |
|---|---|
| IPR001356:Homeobox | 3.98E-19 |
| IPR012287:Homeodomain-related | 1.30E-18 |
| IPR017970:Homeobox, conserved site | 6.36E-17 |
| IPR013087:Zinc finger, C2H2-type/integrase, DNA-binding | 2.99E-05 |
| IPR000910:High mobility group, HMG1/HMG2 | 2.45E-04 |
| IPR007087:Zinc finger, C2H2-type | 2.47E-04 |
| IPR015880:Zinc finger, C2H2-like | 3.85E-04 |
| IPR003654:Paired-like homeodomain protein, OAR | 4.24E-03 |
| IPR001092:Basic helix-loop-helix dimerisation region bHLH | 7.62E-03 |
| IPR013088:Zinc finger, NHR/GATA-type | 8.85E-03 |
| IPR000047:Helix-turn-helix motif, lambda-like repressor | 2.37E-02 |
| IPR012682:Transcription regulator Myc, N-terminal | 2.53E-02 |
| IPR002418:Transcription regulator Myc | 2.53E-02 |
| IPR004979:Transcription factor AP-2 | 3.77E-02 |
| IPR013854:Transcription factor AP-2, C-terminal | 3.77E-02 |

**Table 5-7. Significantly detected GO-BP in upregulated transcription factors during 1st ZGA (FDR < 0.05)**

| GO term | Count | % | FDR |
|---|---|---|---|
| GO:0006355~regulation of transcription, DNA-templated | 45 | 22.4 | 3.41E-23 |
| GO:0006351~transcription, DNA-templated | 44 | 21.9 | 9.78E-21 |
| GO:0045944~positive regulation of transcription from RNA polymerase II promoter | 33 | 16.4 | 5.53E-11 |
| GO:0006357~regulation of transcription from RNA polymerase II promoter | 23 | 11.4 | 4.10E-10 |
| GO:0006366~transcription from RNA polymerase II promoter | 17 | 8.5 | 4.04E-09 |
| GO:0000122~negative regulation of transcription from RNA polymerase II promoter | 24 | 11.9 | 1.24E-07 |
| GO:0008285~negative regulation of cell proliferation | 13 | 6.5 | 1.05E-02 |
| GO:0045892~negative regulation of transcription, DNA-templated | 13 | 6.5 | 1.67E-02 |
| GO:0001709~cell fate determination | 5 | 2.5 | 3.42E-02 |
| GO:0006338~chromatin remodeling | 7 | 3.5 | 3.79E-02 |
| GO:0060021~palate development | 8 | 4.0 | 4.87E-02 |

**Table 5-8. Significantly detected GO-BP in upregulated transcription factors during 2nd ZGA (FDR < 0.05)**

| GO term | Count | % | FDR |
|---|---|---|---|
| GO:0006355~regulation of transcription, DNA-templated | 36 | 31.6 | 1.30E-22 |
| GO:0006351~transcription, DNA-templated | 31 | 27.2 | 8.02E-16 |
| GO:0007275~multicellular organism development | 18 | 15.8 | 1.99E-09 |
| GO:0006366~transcription from RNA polymerase II promoter | 14 | 12.3 | 1.09E-08 |
| GO:0045944~positive regulation of transcription from RNA polymerase II promoter | 22 | 19.3 | 2.83E-07 |
| GO:0000122~negative regulation of transcription from RNA polymerase II promoter | 16 | 14.0 | 8.79E-05 |
| GO:0009952~anterior/posterior pattern specification | 9 | 7.9 | 1.04E-04 |
| GO:0045665~negative regulation of neuron differentiation | 7 | 6.1 | 1.60E-03 |
| GO:0045892~negative regulation of transcription, DNA-templated | 11 | 9.6 | 5.01E-03 |
| GO:0048863~stem cell differentiation | 5 | 4.4 | 2.29E-02 |
| GO:0060021~palate development | 7 | 6.1 | 2.44E-02 |
| GO:0045893~positive regulation of transcription, DNA-templated | 11 | 9.6 | 2.59E-02 |

**Table 5-9. Primers used for the exon-intron RT-PCR**

| Gene name | Wave | Location | Primer sequences | |
|---|---|---|---|---|
| | | | Forward (5'→3') | Reverse (5'→3') |
| WNT3A | 2nd | Exon 3 - Intron 3 | CTTTTGCAGTGACCAGG TCC | TGCTGCCTGTTTTGATC CAC |
| C8ORF22 | 2nd | Exon 3 - Intron 3 | CACCGACTGTTGGCAGA AAA | ACAGATTCAAAAAGGT TAGGAGTAGG |

## 4. Discussion

Induced pluripotent or totipotent stem cells by the defined factors have wider applications such as regenerative medicine and production of new disease models in mammals. It was accomplished based on the comprehensive understanding transcriptional regulation and the related factors in early development. Reprogramming of somatic cells in avian species could be utilized as the genetic resources for restoration of endangered birds, research of avian stem cells, and biotechnology. However, the approaches to date has been limited in not fully reprogrammed cells depending on mammalian factors (Lu et al., 2012, Rossello et al., 2013, Lu et al., 2014, Choi et al., 2016, Kim et al., 2017). Thus, exploring the mechanism of early developmental stages is crucial for establishing higher potent-states.

In mouse, minor ZGA occurs in the pronuclear stages, followed by transcriptional silencing until major ZGA at the 2-cell stage (Aoki et al., 1997, Braude et al., 1988, Xue et al., 2013). In chickens, we suggest that the first wave of transcriptional activation occur between the oocyte and zygote stages. The term “chromatin remodeling” was also associated with the first wave, which is consistent with the chromatin changes that occur in mammals after fertilization and prior to the onset of minor ZGA (Aoshima et al., 2015, Wu et al., 2016).

MuERV-L derived from endogenous retroviruses is expressed by minor ZGA in totipotent 2C (Macfarlan et al., 2012). Similarly, we observed InterPro domains associated with oncogenes, such as Myb-type HTH DNA-binding domains and Ets which are derived from chicken retroviruses E26, were induced during first wave. TFs related to Ets and Myb-type HTH DNA-binding domains seem to be related to totipotency or cell proliferation in early

development in chickens, since these families are proto-oncogenes and Myb-Ets fusion protein can transform hematopoietic progenitor cells (Metz and Graf, 1991, Graf et al., 1992). On the other hand, chicken ESCs-specific endogenous retrovirus *ENS-1* (Acloque et al., 2001, Mey et al., 2012) was not induced by first wave.

Among the mammalian factors used in avian reprogramming, *POUV* and *SOX2* were involved only in the second transition (from EGK.III to EGK.VI). Even after the second wave, *SOX2* was expressed at low levels throughout intrauterine development. In other species, maternally inherited TFs, including *NANOG*, *POUV*, and *SOX2* homologues, are associated with ZGA (Lee et al., 2013, Lee et al., 2014). Among the three homologues in chicken, *NANOG* alone was inherited maternally and was up-regulated in both waves. However, *POUV* and *SOX2* were not significantly up-regulated between the oocyte and zygote stages. Thus, other members of the POU or SOX family may be involved in chicken ZGA because of the conserved mechanisms in other vertebrates. *LIN28A* was also not involved in first wave. Conversely, other reprogramming factors, MYC and KLF4, were induced by both waves. In addition, *MYC* that is involved in ZGA in frogs and mice (Gusse et al., 1989, Zeng and Schultz, 2005), was inherited maternally and induced after both waves. Because *MYC* initializes reprogramming in frogs and mice, it may also be involved in initiating second wave in chickens.

In contrast, high-mobility group proteins, HMG1/HMG2, were found exclusively during the area pellucida formation period. Furthermore, *TCF7L2* and *LEF1* TFs (Cadigan and Waterman, 2012), major mediators of Wnt signalling, were up-regulated only after the second wave. The nuclear localization of β-catenin, an important component of canonical Wnt

signalling, was first observed in EGK.VIII at the developing area opaca and marginal zone (Roeser et al., 1999). During area pellucida formation, the anterior-posterior axis is initially formed when gravitational forces reorient the yolk and determine the posterior side (Kochav and Eyal-Giladi, 1971). Thus, Wnt signalling may be the key regulator of axis formation in chicken early development. Additionally, TFs related to anterior/posterior pattern specification were enriched after the second wave. Furthermore, many Hox family genes associated with the term “stem cell differentiation” were expressed concurrently with morphological differentiation during area pellucida formation. These results indicate that after second wave, cellular states would be getting lower potency compared to cleavage stages.

This was supported by the previous results that lineage specification already occurs following the salt-and-pepper manner before morphological segregation (Kochav and Eyal-Giladi, 1971, Mak et al., 2015). Moreover, chicken ESCs derived from EGK.VI to X could not generate germline chimera (Pain et al., 1996, van de Lavoir et al., 2006). In avian species, primordial germ cells (PGCs) are feasible for conservation of endangered birds and transgenesis, because the unique physiologies of avian embryonic development disable the methods in mammals (Han, 2009). The specification of PGCs seems to be predetermined by germ plasm (Tsunekawa et al., 2000, Lee et al., 2016). However, the precursors of PGC were merely observed in cleavage stage and PGCs are finally specified after second wave of gene expression (Lee et al., 2016). According to this mechanism, germline-potent cells could be generated from the reprogrammed cells in totipotent state combined with germ plasm factors.

Taken together, our results explore the transcriptional regulation and

related TFs during early chicken development and are compared with mammals. Also, based on the examples in mammals and the limitation in birds, our study provides new insight into the reprogramming strategy in avian species. Our findings contribute to large number of reprogrammable factors and suggest that the reprogramming in birds pursue the mimicking higher-potent cleavage stage.

# CHAPTER 6

# Avian Zygotes Activate Only Maternal Alleles to Inhibit Variation due to Supernumerary Sperm

## 1. Introduction

In animals, the highly coordinated genetic events of early embryogenesis are governed by the massive induction of transcripts, called zygotic genome activation (ZGA) (Lee et al., 2014). The massive transcriptional activation through the ZGA governs the establishing of individual organism. In mammals, the first-wave of transcriptional activation occurs after fertilization during pronucleus (PN) formation (Lee et al., 2014). During the pronuclear stage, the most distinctive feature is that transcriptional activation of the paternal PN is greater than that of maternal PN due to its chromatin state and allelic histone modifications (Bouniol et al., 1995, Aoki et al., 1997, Wu et al., 2016, Aoshima et al., 2015, Zhang et al., 2016). As well as epigenetic regulation, transcriptional activation is also highly promiscuous, such as intergenic region and untranslatable mRNAs in 1-cell stage in mouse (Abe et al., 2015). In zebrafish, 1-cell embryo activates the mitochondrial genome (Heyn et al., 2014). Moreover, the earliest genes of transcriptional activation were evolutionarily varied as in the comparison among the species (Heyn et al., 2014). Even, distinct transcriptional activities by the parental genomes were shown in plant zygotes (Anderson et al., 2017). Thus, the mechanism of transcriptional activation need to be examined in each species.

Compared to mammals, polyspermy is indispensable for initiating of early embryonic development in bird without any pathological defects (Snook et al., 2011). Multiple sperm result in long-lasting small $Ca^{2+}$ rise, which finally induce egg activation (Iwao, 2012). Despite the crucial period for the initial stages of development, there have been no detailed investigations regarding the dynamic events of transcriptional activation at the moment of fertilization in avian species using transcriptomic approaches.

Here, to explore the transcriptional characteristics during chicken embryogenesis, we performed whole-transcriptome sequencing (WTS) throughout chicken early development from oocyte to Eyal–Giladi and Kochav Stage X (EGK.X). Firstly, our genome-wide study of primary transcripts, traced by intronic mapped reads, clarified that the definite transcriptional activation during embryogenesis in chicken. Moreover, since polyspermic fertilization in avian species contrary to mammals, we traced allelic expression during 1$^{st}$ wave of transcriptional activation. To investigate which parental genome activated, we conducted additional multi-omics approaches, including whole-genome sequencing (WGS) of two different parental breeds to identify breed-specific single nucleotide polymorphisms, and WTS of their single hybrid embryo to determine allelic expression upon transcriptional activation. Our results firstly provide intriguing results about the first genome activation associated with the physiological characteristic upon fertilization in birds.

## 2. Materials and methods

### *Experimental animals and animal care*

The management and experimental use of chickens were approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1). The experimental animals were cared according to a standard management programme at the University Animal Farm, Seoul National University, Korea. The procedures for animal management, reproduction, and embryo manipulation adhered to the standard operating protocols of our laboratory.

### *Genomic DNA isolation and DNA sequencing library preparation for whole genome sequencing (WGS)*

Genomic DNA was isolated from blood collected from the wing vein of 6 parental chickens [3 male Korean Oge (mKO) and 3 female White Leghorn (fWL)], using 1-mL 30-gauge syringes (Shina Corporation, Seoul, Korea). The collected blood samples were transferred into EDTA tubes (BD Biosciences, San Jose, CA, USA) immediately after collection. 10 μL of blood were used for isolation of genomic DNA using a DNeasy Mini Kit (Qiagen, Valencia, CA, USA). The quality of the extracted genomic DNA was determined using the Trinean DropSense96 system (Trinean, Gentbrugge, Belgium), RiboGreen (Invitrogen, Carlsbad, CA, USA), and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Genomic DNA was used for construction of cDNA libraries using a TruSeq Nano DNA LT Library Preparation Kit (Illumina Inc., San Diego, CA, USA). The resulting libraries were subjected to chicken genome resequencing (30× coverage) using the Illumina Nextseq 500 platform to produce paired 150-bp reads. The

raw sequencing data were deposited in the NCBI SRA database (Project #: PRJNA393895).

***RT-PCR for confirmation of hybrid embryos***

Before collecting early embryos, Eyal-Giladi and Kochav stage(Eyal-Giladi and Kochav, 1976) X (EGK.X) blastoderms between mKO and fWL were incubated in a chamber at 37.5°C under 80% humidity for 18 h. Genomic DNA was isolated from Hamburger and Hamilton stage (Hamburger and Hamilton, 1951) 4 (HH4) embryos using a DNeasy® Mini Kit (Qiagen). RT-PCR was performed to confirm hybridisation between KO and WL using breed-specific primers (AS3554-I9 / P5FWD WL F: 5'-AGC AGC GGC GAT GAG CGG TG-3'; WL R: 5'-CTG CCT CAA CGT CTC GTT GGC-3'; AS3554-WT / P5FWD KO F: 5'-AGC AGC GGC GAT GAG CAG CA-3'; KO R: 5'-CTG CCT CAA CGT CTC GTT GGC-3') (Choi et al., 2007), with an initial incubation at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, 69°C for 30 s, and 72°C for 30 s. The reaction was terminated after a final incubation at 72°C for 10 min.

***Chicken early hybrid embryo preparation, RNA isolation, and RNA-Seq library preparation for whole-transcriptome sequencing (WTS)***

The egg-laying times of three fWL, which were mated with mKOs, were recorded. A single hybrid EGK.X blastoderm was collected from WL hens after oviposition. To collect single oocytes and hybrid zygotes, WL hens were sacrificed and the follicles were collected. Oocytes and hybrid zygotes were collected simultaneously from one WL hen. Due to the small transcriptomic differences between pre- and post-ovulatory oocytes in the previous study(Elis et al., 2008) and the infeasibility of simultaneous

acquisition of post-ovulatory oocytes and zygotes from a single hen, we isolated only the pre-ovulatory large F1 oocyte. Only zygote embryos not showing cleavage and located in the magnum were collected within 1 h after fertilisation according to the recorded egg-laying times. All embryos were classified according to morphological criteria (Figure 6-1A). Shortly after collection, the embryos were separated from the egg using sterile paper, and the shell membrane and albumen were detached from the yolk. A piece of filter paper (Whatman, Maidstone, UK) with a hole in the centre was placed over the germinal disc. After cutting around the paper containing the embryo, it was gently turned over and transferred to saline to further remove the yolk and vitelline membrane to allow embryo collection. Total RNA was isolated from early embryos using TRIzol reagent (Invitrogen). The quality of the extracted total RNA was determined using the Trinean DropSense96 system (Trinean), RiboGreen (Invitrogen), and an Agilent 2100 Bioanalyzer (Agilent Technologies). Total RNA was used for construction of cDNA libraries using a TruSeq Stranded Total RNA Sample Preparation Kit (Illumina, Inc.). The resulting libraries were subjected to whole-transcriptome analysis using the Illumina Nextseq 500 platform to produce paired 150-bp reads. The raw sequencing data were deposited in the GEO database (GSE100798).

***Alignment and variant calling for WGS data***

The paired-end reads for 6 chickens (3 biological replications in mKO and fWL breeds) were generated using the Illumina Nextseq 500 platform. In total, 8.38 billion reads or ~2.53 Gbp of sequences were generated. For generating clean reads, paired-reads sequences were qualitied using Trimmomatic (v0.33) (Bolger et al., 2014). Using Bowtie 2 (v2.2.5) (Langmead and Salzberg, 2012), reads were aligned to the reference genome

sequence galGal4 (Build v 4.82) with an average alignment rate of 91.61%. After filtering the potential PCR duplicates and correcting for misalignments due to the presence of insertions and deletions (INDELs), we detected single nucleotide polymorphisms (SNPs) using GATK v3.4.46(McKenna et al., 2010). More detailed, potential PCR duplicates were filtered using the option "REMOVE_DUPLICATES = true" in "MarkDuplicates" open-source tool of Picard (v 1.138) (https://broadinstitute.github.io/picard/). After that, SAMtools (v1.2)(Li et al., 2009) was employed to create index files for reference and Binary Alignment/Map (BAM) files. In variant calling step with GATK v3.1, local realignment of reads to correct misalignments was performed because the presence of INDELs ("*RealignerTargetCreator*" and "*IndelRealigner*" arguments). In the GATK tool, two types of arguments, "*UnifiedGenotyper*" and "*SelectVariants*" were employed for variant calling. In addition, "*VariantFiltration*" was applied to filter bad variants based on the following criteria: (1) Variants with a phred-scaled quality score < 30 were filtered; (2) SNPs with "*mapping quality zero (MQ0) > 4*", "*quality depth < 5*", and "*(MQ0 / (1.0*DP)) > 0.1*" were filtered; and (3) SNPs with "*phred-scaled P value using Fisher's exact test > 200*" were filtered, which are indicative of false-positive calls. As a result, 10,529,469 variants were detected and 9,805,997 variants (93.129%) were known variants previously (Table 6-8).

***Quality control, alignment, and quantification of mapped reads for RNA-Seq data***

To generate clean reads, Trimmomatic (v 0.33) (Bolger et al., 2014) was used. Per-base sequence qualities were checked using FastQC (v 0.11.2) (Andrews, 2010) and filtered fastq files. Trimmed reads were aligned to the

galGal4 genome files using the HISAT2 alignment software (v 2.0.0) (Kim et al., 2015) with the following alignment option: "*--rna-strandness RF*". Sequence Alignment/Map (SAM) files were converted into compressed and sorted BAM files using SAMtools (v 1.4.1) (Li et al., 2009). We quantified the mapped reads using HTSeq-count (Anders et al., 2014) with the merged gene annotation file (.GTF), with total RNAs and lincRNAs derived from Ensembl and ALDB (Li et al., 2015), respectively. In the quantification step, four types of genomic area were considered such as transcripts, exons, intron, and intergenic regions. Although quantification of the transcript and exon's level is directly available without any pre-processing step by using galGal4 GTF file, genomic position should be defined for estimating expression levels of the intron and intergenic regions. When defining intron area, the overlapped annotation of the exon within associated gene makes it hard to define intron regions on the reference genome. In addition, different stranded information should be considered when defining intron regions between each exon. For considering these issues, we defined intron region using custom python script and implemented following Pseudo code 1:

| **[Pseudo code1] Defining intron region based on the Ensembl GTF file (galGal4)** |
| --- |
| First, exon combining step should be performed in order to consider exon-overlap pattern as follows:<br><br>For *i* in 1 to (# of genes):<br><br>For j in 2 to (# of Exons within associated gene - 1): |

Let $E_1$ and $E_2$ be the (*j-1*)th and *j*-th exons, respectively.

$$CE_S{:}\,CE_E = \begin{cases} \boldsymbol{E_{1,S}{:}\,E_{2,E}} & if\ E_{1,S} \le E_{2,S}\ and\ E_{1,E} \le E_{1,E} \le E_{2,E} \quad \text{(Case 1)} \\ \boldsymbol{E_{2,S}{:}\,E_{1,E}} & if\ E_{2,S} \le E_{1,S} \le E_{2,E}\ and\ E_{2,S} \le E_{1,E} \quad \text{(Case 2)} \\ \boldsymbol{E_{1,S}{:}\,E_{1,E}} & if\ E_{1,S} \le E_{2,S}\ and\ E_{2,E} \le E_{1,E} \quad \text{(Case 3)} \\ \boldsymbol{E_{2,S}{:}\,E_{2,E}} & if\ E_{2,S} \le E_{1,S}\ and\ E_{1,E} \le E_{2,E} \quad \text{(Case 4)} \end{cases}$$

where $CE_S{:}\,CE_E$ represents start and end positions of the combined exon considering overlapped pattern, $E_1$ and $E_2$ represents 1st and 2nd exons within the gene based on the position of the reference genome, and sub-notation of the *S* and *E* represents start and end positions. There are 4 overlap patterns between exon annotation within the gene such as:

**(Case1):** Partial exon overlap between 3' end of the $E_1$ and 5' of the $E_2$.

**(Case2):** Partial exon overlap between 3' end of the $E_2$ and 5' of the $E_1$, which means gene is annotated in opposite strand.

**(Case3):** Complete exon overlapped, $E_1$ is annotated and located within the $E_2$.

**(Case4):** Complete exon overlapped, $E_2$ is annotated and located within the $E_1$, which means gene is annotated in opposite strand.

Based on calculated $CE_S{:}\,CE_E$ list in each gene, intron regions can be

easily defined as follows:

For *i* in 1 to (# of genes):

For *j* in 2 to (# of $CE_{iS}$: $CE_{iE}$ within the associated gene):

$$I_{i,j,s}: I_{i,j,E} = (CE_S: CE_E[\mathrm{j}-1)]_E + 1 : (CE_S: CE_E[\mathrm{j})]_S - 1$$

where $CE_{iS}$: $CE_{iE}$ is the list of combined exon for *i* th gene and $I_{i,j,s}$: $I_{i,j,E}$ represents coordinates of the start:end site for intron regions.

Like as the method of calculating the intron region between exons within the associated gene, the intergenic region among the genes within the same chromosome can be defined as following Pseudo code2:

**[Pseudo code2] Defining intergenic region based on the Ensembl GTF file (galGal4)**

First, combining step of the gene annotation within same chromosome should be performed for considering gene overlap pattern as follows:

For *i* in 1 to (# of chromosomes):

For j in 2 to (# of Genes within associated chromosome - 1):

Let $G_1$ and $G_2$ be the (*j-1*)th and *j*-th genes,

respectively.

$$CG_S{:}\,CG_E = \begin{cases} \boldsymbol{G_{1,S}{:}\,G_{2,E}} & if\ G_{1,S} \le G_{2,S}\ and\ G_{1,E} \le G_{1,E} \le G_{2,E} \quad \text{(Case 1)} \\ \boldsymbol{G_{2,S}{:}\,G_{1,E}} & if\ G_{2,S} \le G_{1,S} \le G_{2,E}\ and\ G_{2,S} \le G_{1,E} \quad \text{(Case 2)} \\ \boldsymbol{G_{1,S}{:}\,G_{1,E}} & if\ G_{1,S} \le G_{2,S}\ and\ G_{2,E} \le G_{1,E} \quad \text{(Case 3)} \\ \boldsymbol{G_{2,S}{:}\,G_{2,E}} & if\ G_{2,S} \le G_{1,S}\ and\ G_{1,E} \le G_{2,E} \quad \text{(Case 4)} \end{cases}$$

where $CG_S{:}\,CG_E$ represents start and end positions of the combined gene considering overlapped pattern, $G_1$ and $G_2$ represents 1st and 2nd genes within the same chromosome based on the position of the reference genome, and sub-notation of the *S* and *E* represents start and end positions. There are 4 overlap patterns between gene annotation within the same chromosome such as:

**(Case1):** Partial gene overlap between 3' end of the $G_1$ and 5' of the $G_2$.

**(Case2):** Partial gene overlap between 3' end of the $G_2$ and 5' of the $G_1$, which means gene is annotated in opposite strand.

**(Case3):** Complete gene overlapped, $G_1$ is annotated and located within the $G_2$.

**(Case4):** Complete gene overlapped, $G_2$ is annotated and located within the $G_1$, which means gene is annotated in opposite strand.

Based on calculated $CG_S{:}\,CG_E$ list in each chromosome, intergenic

regions can be easily defined as follows:

For $i$ in 1 to (# of chromosomes):

For $j$ in 2 to (# of $CG_{iS}$:$CG_{iE}$ within the associated chromosome):

$$T_{i,j,s}:T_{i,j,E} = (CG_S:CG_E[\mathrm{j}-1)]_E + 1 : (CG_S:CG_E[\mathrm{j}])]_S - 1$$

where $CG_{iS}$:$CG_{iE}$ is the list of combined exon for $i$ th gene and $T_{i,j,s}$:$T_{i,j,E}$ represents coordinates of the start:end site for intergenic regions.

After defining intronic and intergenic regions, gene annotation file (.GTF form) was generated based on the coordinate information. Based on these GTF files, expression levels were measured with HTseq-count (v 0.6.1) (Anders et al., 2015).

***Variant calling RNA-Seq***

Using the alignment file (.BAM), potential PCR duplicates were removed using the Picard (v 1.138) with "*REMOVE_DUPLICATES = true*" in "*MarkDuplicates*" option. After that, SplitNCigarReads tool implemented in GATK was performed with "*-rf ReassignOneMappingQuality -RMQF 255 -RMQT 60 -U ALLOW_N_CIGAR_READS*" option. In variant calling step with GATK, local realignment of reads was performed to correct misalignments ("*RealignerTargetCreator*" and "*IndelRealigner*" options). Finally, base-recalibration was performed using BaseRecalibrator implemented in GATK

with known variant sites in galGal4. Using HaplotypeCaller in the GATK tool, variant calling was performed with "*-dontUseSoftClippedBases -stand_call_conf 20.0 -stand_emit_conf 20.0*" option. Finally, bad variants were filtered using the VariantFiltration tool with "*-window 35 -cluster 3 -filterName FS -filter "FS > 30.0" -filterName QD -filter "QD < 2.0" option*. As a result, 265,788 variants were detected and 248,030 variants (93.319%) were previously known sites (Table 6-9).

***Identification of the maternal and paternal expressed genes through breed specific variants detection***

Maternal and paternal samples were genotyped using WGS and their offspring including maternal oocyte were genotyped using WTS (variant calling on the RNA-Seq data). After pre-processing, there are two genotype data (DNA and RNA sequencing data) across the mother, father, oocyte, zygote, and EGK.X. In two types of SNP data, 10,529,469 and 265,788 variants were detected in DNA and RNA sequencing data, respectively. First, breed-specific SNPs such as: (1) SNPs "0/0" and "1/1" genotype for maternal and paternal groups, respectively; (2) SNPs "1/1" and "0/0" genotype for maternal and paternal groups, respectively, were identified and annotated using SnpSift (Cingolani et al., 2012) in parental SNP data. As a result, 216,003 SNPs were identified as breed-specific SNPs. After that, two SNP datasets (breed-specific SNP and their offspring genotype derived from the RNA-Seq data) were combined to detect maternal and paternal expressed genes and 14,817 SNPs were commonly identified in breed-specific SNP and RNA-Seq data derived SNP. Using this combined genotype data, three types of filtering steps were carried out. First, mismatched genotype of the reference and alternative allele between breed-specific SNP and SNPs derived from the

RNA-Seq were removed and two variants were removed in this step. Second, different genotypes within the biological replications were removed and 9,143 SNPs were removed this step. Finally, mismatched genotype between maternal samples and oocyte samples were removed and 6 SNPs were removed. Finally, 5,666 SNPs were left and they were annotated by SnpSift tool with galGal4 and ALDB (.GTF) files. In order to find the most conservative evidence of the parental expression, we filtered only a single SNP was found within the gene or genotypes pattern is not consistent among the SNPs were finally filtered out. In addition, unannotated SNPs in both databases, Ensembl and ALDB, were removed for biological interpretation. As a result, 1,544 SNPs were finally detected as parental expression markers, all of which showed a maternal expression pattern.

***Identification of the differentially expressed regions such as exon, transcript, intron, and intergenic between early developmental stages of the chicken***

In the multi-omics analysis for identifying phenomenon of the maternal and paternal expression, RNAs derived from the three sources such as oocyte, zygote, and EGK.X were employed and sequenced. For exploring gene expression change during intermediate stages, we employed pre-existing WTS data including oocyte, zygote, EGK.I, EGK.III, EGE.VI, EGK.VIII, and EGK.X stages (GSE86592). We quantified mapped reads of the intron and intergenic regions using these data. As a result, three types of matrix data were generated and these data were employed in statistical analysis. Total six statistical tests, oocyte vs. zygote, zygote vs. EGK.I, EGK.I vs. EGK.III, EGK.III vs. EGK.VI, EGK.VI vs. EGK.VIII, and EGK.VIII vs. EGK.X, were performed using the edgeR package (Robinson et al., 2010), in the matrix data

derived from intron and intergenic regions, respectively. More detailed, contrast tests were performed on the one-way analysis of deviance model (one-way ANODEV), as follows:

$$\log(E(Expression_i)) = \mu + Stage_i \qquad \text{(Eq. 1)}$$

,where the *Stage* represents each pair of the developmental stages, and *i* is subjects. Significant result was considered at a FDR adjusted $P < 0.05$ (Benjamini and Hochberg, 1995).

***Identification of functional characteristics of the detected DEGs***

Based on the biological process terms (BP-term) of the Gene Ontology (GO) and KEGG pathways, functional enrichment test was performed with significantly detected DEGs. In addition, transcription factors (TFs) (AnimalTFDB; http:// http://www.bioguo.org/AnimalTFDB) and SMART protein domain databases (Letunic et al., 2011) were also employed to characterise observed genes.

***Exon–intron RT-PCR and validation of allelic expression***

Total RNA (1 µg) was used as the template for cDNA synthesis using the SuperScript III First-Strand Synthesis System (Invitrogen). The cDNA was serially diluted five-fold and equalised quantitatively for PCR amplification. To validate allelic expression, additional single hybrid embryos at EGK.III and VI were collected from parents with genotypes identical to WGS, and their total RNA isolation and cDNA synthesis were performed as described above. Primers for exon–intron PCR of six genes and for allelic expression of six genes were designed using the programme Primer3

(Untergasser et al., 2012) (Table 6-10 and 6-11). RT-PCR was performed with an initial incubation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 59°C for 30 s, and 72°C for 30 s. The reaction was terminated after a final incubation at 72°C for 5 min. PCR products were cloned into the pGEM-T Easy Vector (Promega, Madison, WI, USA) for sequencing with an ABI 3730xl DNA Analyzer (Applied Biosystems).

## 3. Results

### *Genome-wide analysis of transcriptional activation in chicken early development*

To explore zygotic transcription during chicken early development, we performed WTS of pre-ovipositional embryos from the oocyte stage to EGK.X (Figure 6-1A). Unsupervised hierarchical clustering of transcriptome profiles showed that zygote, similar to EGK.I and EGK.III embryos, differed from the oocyte, suggesting dynamic changes in transcriptomic feature after fertilization in chicken (Figure 6-1B). Cells positive for the phosphorylation of RNA polymerase II C-terminal domain first appeared during the late EGK.II to early EGK.III stage (Nagai et al., 2015); however, actual transcriptional changes were found between EGK.III and EGK.VI at the transcriptome level (Figure 6-1B), indicating $1^{st}$ and $2^{nd}$ wave of transcriptional activation in chicken comparable to the events of other species.

To deeply understand zygotic transcription, genome-wide analysis of the number of expressed genes and their regional distributions during early development in the chicken were investigated. Quantitatively, the number of expressed regions (including exon, intron, and intergenic regions) across whole chicken genome were significantly different between oocyte and zygote and between EGK.III and EGK.VI (Figure 6-2), which indicated the large-scale degradation of stored transcripts in the oocyte after fertilization and increasing transcript number after EGK.VI. Next, we analyzed the proportion of mapped reads throughout early development (Figure 6-3). Like Figure 6-3, the proportion of expressed intronic regions were reduced after fertilization and increased gradually after EGK.VI. Unlike minor ZGA in mammals (Abe et al., 2015), the proportion of expressed intergenic regions of reads were

similar during pre-ovipositional development in chicken regardless of the transcriptional changes.

In case of expressed genic region, differentially expressed genes (DEGs) were detected during each consecutive stage, even between the oocyte and zygote, and classified in messenger RNA (mRNA), microRNA precursor (pre-miR), small nucleolar RNA (snoRNA), miscellaneous RNA (miscRNA), and small nuclear RNA (snRNA) expression from Ensembl gene annotations and long intergenic non-coding RNA (lincRNA) annotations from The domestic-animal lncRNA database (ALDB) database (Li et al., 2015) (false discovery rate [FDR] adjusted $P < 0.05$) (Figure 6-4). The mRNA and lincRNA expression patterns were similar in terms of the number of detected DEGs. Specifically, large numbers of up- and down-regulated genes were observed in the 1st wave, while other RNAs were mostly down-regulated after fertilization, suggesting the potential role of long transcripts in early cleavage stages. Conversely, all RNAs with significant differential expression were up-regulated in the 2nd wave, suggesting maternal-to-zygotic transition (MZT) after EGK.VI. We also investigated primary transcripts throughout early development to obtain definitive evidence of gene activation as the massive alteration of maternally stored RNAs in DEGs after fertilization may have masked the smaller effects of the 1st activation. The results indicated up-regulation of intronic-span mapped reads of each gene between oocyte and zygote, demonstrating nascent transcription by the 1st wave, as well as between EGK.III and EGK.VI by the 2nd wave (Figure 6-1C).

***The detailed investigation of activated genes defines 1st and 2nd wave of transcriptional activation in chicken.***

To confirm and validate the activation of transcription, we examined

candidate genes based on intronic expression during the two waves of activation at base resolution (Table 6-1). Genome viewers in Figure 6-5 clearly show two waves of transcriptional activation between oocyte and zygote and between EGK.III and EGK.VI. In addition, exon–intron RT-PCR of three genes, *distal-less homeobox 6* (*DLX6*), *GATA binding protein 2* (*GATA2*), and *zinc finger protein of the cerebellum 4* (*ZIC4*), in the pre- and post-ovulatory oocyte and zygote showed 1st activation upon fertilization (Figure 6-5A). The same approaches for *wingless-type MMTV integration site family member 11* (*WNT11*), *wingless-type MMTV integration site family member 3A* (*WNT3A*), and *chromosome 2 open reading frame, human C8orf22* (*C8ORF22*) genes indicated a lack of transcriptional activity during rapid cellularization in the cleavage period, and that the 2nd activation of transcription in chicken occurred between EGK.IV and EGK.V, not EGK.II and EGK.III, based on actual transcripts (Figure 6-5B). Consequently, our whole-transcriptome analysis during early development confirmed the definite existence of two waves of transcriptional activation in chicken.

***Multi-omics analysis reveals the exclusive maternal genome activation in 1st wave of transcriptional activation***

Since polyspermic fertilization occurs in avian species, in contrast to mammals, we hypothesized that the haploid nucleus of supernumerary sperms were massively induced for 1st wave of transcriptional activation in addition to paternal and maternal PN activation. To examine this hypothesis, we conducted additional multi-omics approaches using two breeds to investigate which parental genome activated. First, we performed whole-genome sequencing (WGS) of 6 parents [3 male Korean Oge (mKO); 3 female White Leghorn (fWL) chickens] to identify breed-specific single nucleotide

polymorphisms (SNPs) (Figure 6-6A). We also performed WTS of nine single hybrid embryos (biological triplicates per oocyte, zygote, and EGK.X from three pairs of parents) to examine the characteristics of the 1st activated transcripts, including allelic expression during early development in chicken. After confirmation of hybrid embryo formation between mKO and fWL (Figure 6-7), we collected oocyte, zygote, and EGK.X blastoderm from hens on the same day (Figure 6-8). In this study, a single embryo contained an average of 2.1 μg of total RNA (Table 6-2).

In the down-stream analysis, transcriptional relationships among the samples and stages were examined. Triplicate samples of single embryos were clustered into three respective stages based on multidimensional scaling (MDS) (Figure 6-6B). Marked changes in gene expression were observed after fertilization, as the above transcriptome of bulked embryo. A total of 4,275 differentially expressed mRNAs (FDR-adjusted $P < 0.05$) were detected (Figure 6-6*C*), among which 1,883 were up-regulated and 2,392 were down-regulated between the oocyte and zygote stage. In the case of lincRNA expression, 118 were up-regulated and 786 were down-regulated. Because of dramatic changes in early development between fertilization and oviposition, a total of 10,298 mRNAs and 2,507 lincRNAs were designated as DEGs between the zygote and EGK.X stages (Figure 6-6C). In addition, up-regulation of intronic mapped reads was also observed between the single oocyte and zygote stage (Table 6-3; FDR-adjusted $P < 0.05$). These results obtained with single embryos showed a pattern similar to observations made in bulked embryos.

After verification of stable sequencing by single embryo compared to bulked embryos, we then investigated from which parental genome

transcription was activated. To distinguish between maternally and paternally derived transcripts in chicken embryos, we used WGS of parental mKO and fWL and WTS of hybrid single embryos per filtering criteria for detecting allelic expression (See Supplemental Materials and Methods). As a result, 424 parental derived allelic expressions were detected (Table 6-4). Interestingly, all transcripts in the zygote stage showed maternally derived expression even with 1st wave of transcriptional activation (Figure 6-9A). Furthermore, although transcripts by 1st wave were exclusively derived from the maternal alleles, all genes showed biallelic expression and replaced the maternally 1st wave activated genes after MZT in chicken, except for seven mRNAs and two lincRNAs considered to be residual maternal transcripts. Additionally, up-regulated DEGs with intronic expression, including *MAP7 domain containing 1* (*MAP7D1*), *establishment of sister chromatid cohesion N-acetyltransferase 1* (*ESCO1*), *cyclin B3* (*CCNB3*), *synaptotagmin like 1* (*SYTL1*), *grainyhead like transcription factor 1* (*GRHL1*) and *lethal giant larvae homolog 1* (*LLGL1*), induced by the 1st activation (Table 6-5) exclusively involved maternal allelic expression as validated using Sanger sequencing (Figure 6-9B). These genes, except for *GRHL1* earlier at EGK.VI, changed to biallelic expression between EGK.VI and EGK.X. This phenomenon is quite different from mammals with a male pronucleus in which transcriptional activity is two times greater than that of the female pronucleus (Aoki et al., 1997). These observations indicate that there is no possibility of activated transcripts from supernumerary sperm nuclei or paternal PN as in mammals (Bouniol et al., 1995, Aoki et al., 1997). Instead, only the maternal PN was activated after fertilization in avian species.

Functionally, 1st-activated maternal transcripts after fertilization were enriched in the following terms: cell cycle; Notch signaling pathway; Wnt

signaling pathway; regulation of transcription, DNA-templated; and regulation of small GTPase-mediated signal transduction (Figure 6-10A and Table 6-6). To further investigate the characteristics of these terms, we traced the dynamics of gene expression until EGK.X. The 1,883 genes that were up-regulated after fertilization were then categorized per expression pattern between zygote and EGK.X. Three representative patterns were observed: 295 genes did not significantly differ at EGK.X, 221 were significantly up-regulated (FDR < 0.05), and 960 were significantly down-regulated (FDR < 0.05) until EGK.X, respectively (Figure 6-10B). Most transcripts by 1$^{st}$ wave involved in Notch signaling pathway, Wnt signaling pathway, and regulation of small GTPase-mediated signal transduction were down-regulated until EGK.X when MZT was complete. However, expression related to transcriptional activation, including positive regulation of transcription from RNA polymerase II promoter, and regulation of transcription, DNA-templated, was maintained or induced in the 2$^{nd}$ transcription and further development. Along with functional terms, we analyzed transcription factors (TFs) because of the significance of DNA-dependent regulators for downstream gene expression regulating further early development (Zernicka-Goetz et al., 2009, Lee et al., 2014) and the gene sets enriched in DNA-templated regulation of transcription by 1$^{st}$ wave. A total of 115 TFs were up-regulated by 1$^{st}$, and were categorized into three representative patterns (20, 26, and 69 TFs each). Further, the clusters of TFs induced by 1$^{st}$ were classified depending on their roles in chicken early development based on the Simple Modular Architecture Research Tool (SMART) protein domain database (Figure 6-10C and Table 6-7). Specifically, expression of TFs containing helix-loop-helix (HLH), basic region leucine zipper (BRLZ), SWI3, ADA2, N-CoR and TFIIIB (SANT), and Broad-Complex, Tramtrack and Bric a brac (BTB) domains, such as Notch, Wnt, and small GTPase,

seemed to be restricted to the cleavage period. These results suggested that the genetic network regulated by 1st transcription may regulate cellularization and cell polarity during the cleavage period in chicken; however, sustained biallelic gene expression through all intrauterine stages may function in later development.

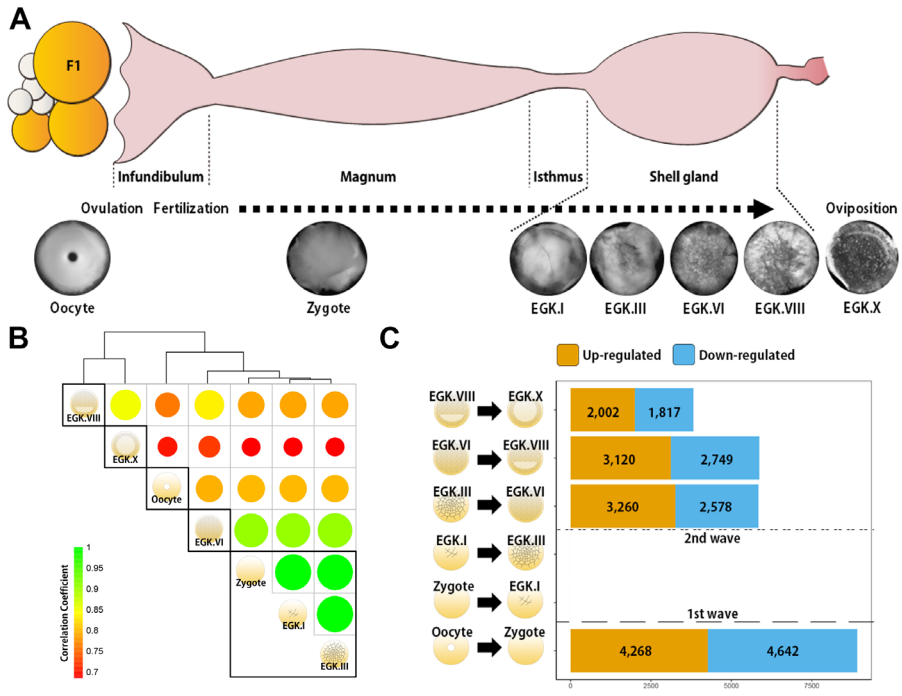


**Figure 6-1. Genome-wide transcriptional activation during chicken early development.** (A) Representative images of early embryos from oocyte to Eyal–Giladi and Kochav X (EGK.X) used for RNA-Seq and acquisition in chicken oviduct. All embryos were classified following the morphological criteria of EGK. (B) Hierarchical clustering of the whole transcriptome during early development in chicken. The size and color of each circle represent strength of the correlation coefficients based on the whole transcriptome expression. The black rectangle represents optimal clusters (k=5) based on the Silhouette score. The transcriptomic changes between consecutive stages, including oocyte vs. zygote and EGK.III vs. EGK.VI, are shown. Zygote, EGK.I, and EGK.III had similar transcriptome profiles. (C) The number of differentially expressed intronic regions between consecutive stages. The Orange and blue color represent Up- and Down-regulated genes at 5% significance level after FDR multiple testing adjustment. The 1$^{st}$ wave between oocyte

and zygote and the 2nd wave of transcriptional activation between EGK.III and EGK.VI were shown.

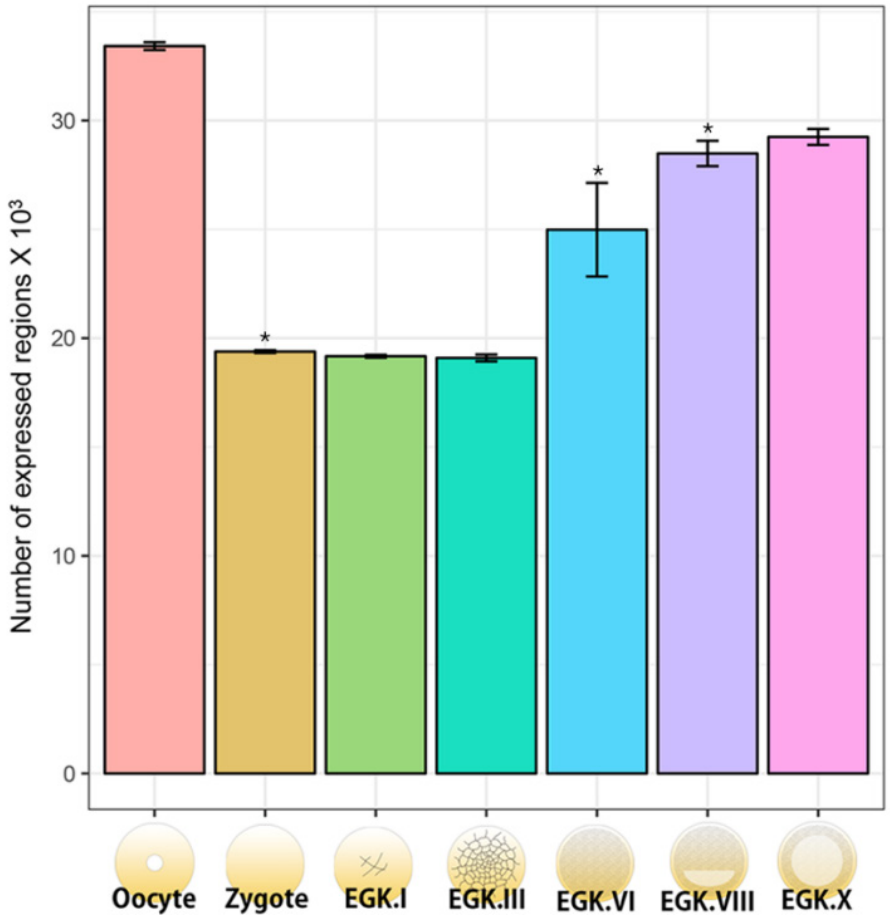


**Figure 6-2. Quantification of the numbers of expressed regions including exon, intron, and intergenic regions on the chicken genome.** The number of expressed regions during chicken early development were investigated based on the quantification result of the mapped reads on the chicken genome (total 188,533 regions were featured). After normalization using trimmed mean of M-value (TMM), the expressed regions were defined based on the numbers of TMM value > 0. Significant differences of the number of annotated regions between consecutive developmental stages (Oocyte vs. Zygote, EGK.III vs. EGK.VI, and EGK.VI vs. EGK.VIII) were represented (pairwise t-test * $P < 0.05$).

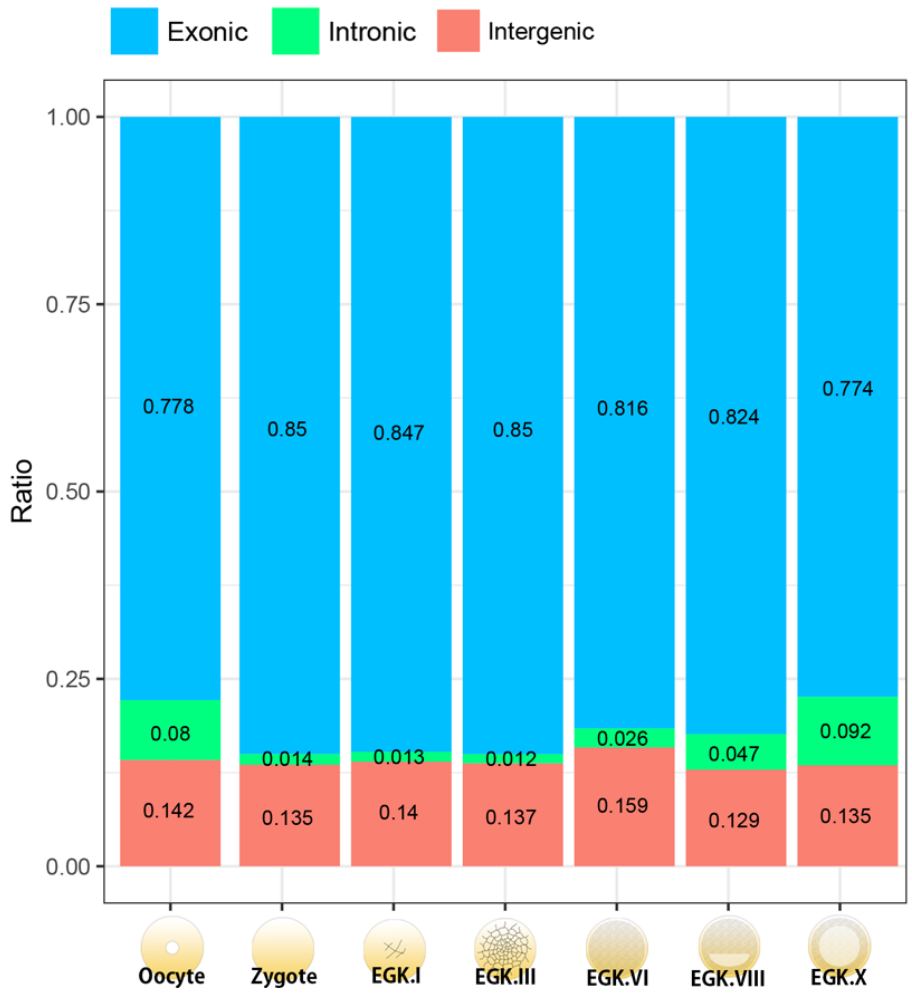


**Figure 6-3. Distribution of mapped reads on the exonic, intronic, and intergenic regions during chicken early development.** The distribution of intronic reads is reduced after fertilization and gradually increased after EGK.VI, probably due to the increase of the exon proportion during both gene activation and processing of maternal RNAs. The proportion of intergenic show little change during pre-ovipositional development in chicken.

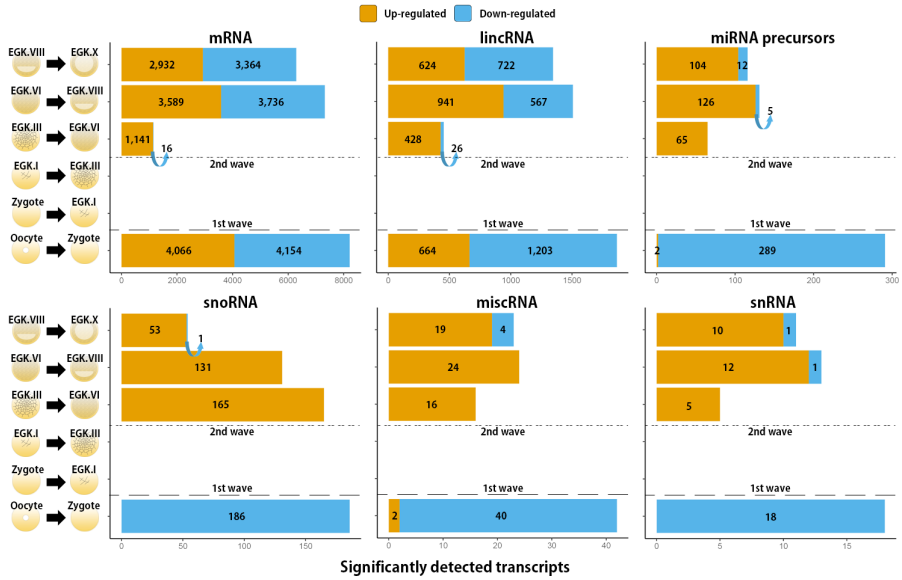


**Figure 6-4. Significantly detected transcripts between each stage during chicken early development.** The number of differentially expressed mRNAs, lincRNAs, miRNA precursors, snoRNAs, miscRNAs, and snRNAs were detected using either Ensembl gene annotation or ALDB database. The Orange and blue color represent Up- and Down-regulated genes at FDR-adjusted $P < 0.05$.

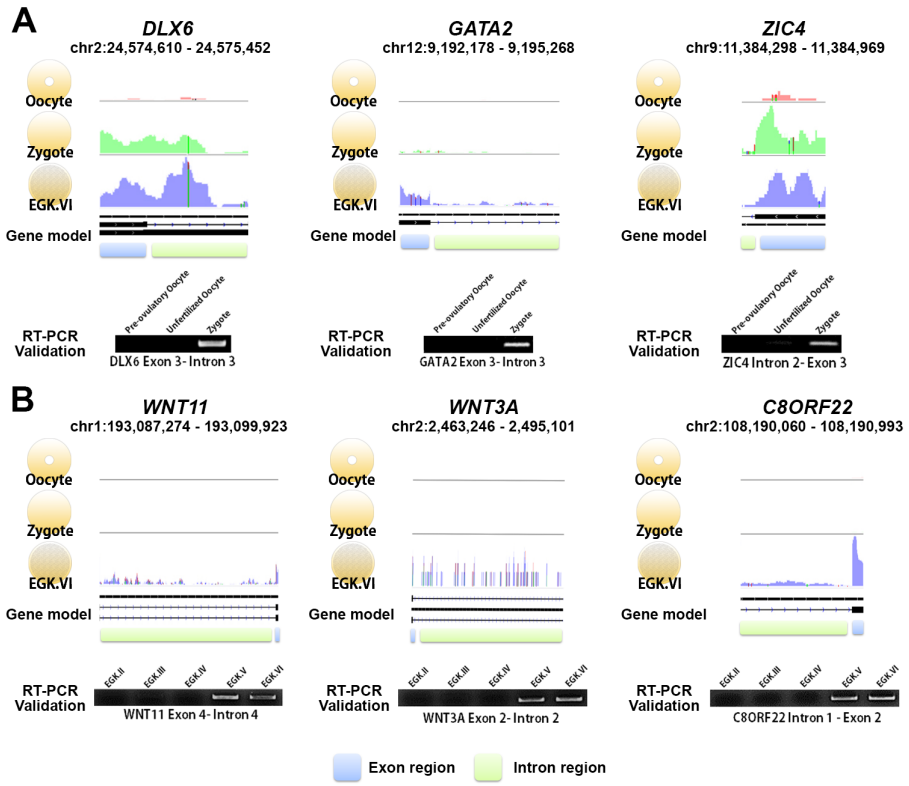


**Figure 6-5. Exonic and intronic mapped reads on candidate genes related to the 1st and 2nd wave of transcriptional activation in chicken.** (A and B) The pooled mapped reads based on the stage (3 samples in each stage) was visualized using Integrative Genomics Viewer tool. Detection of gene activation via the appearance of primary transcripts based on whole-transcriptome sequencing and validation of intronic expression of three genes (*DLX6*, *GATA2*, and *ZIC4*) by 1st wave (A) and other three genes (*WNT11*, *WNT3A*, and *C8ORF22*) by 2nd wave (B) with RT-PCR.

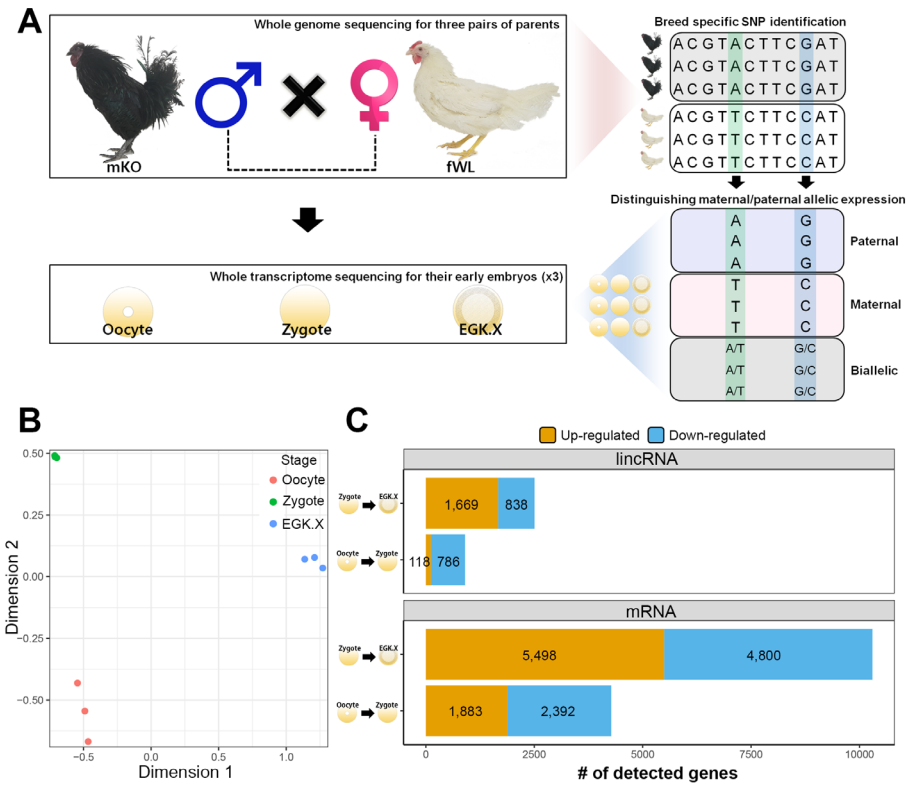


**Figure 6-6. Whole-transcriptome analysis of single early chicken embryos.** (A) Schematic diagram of the experimental design using a multi-omics approach for allelic expression. Three pairs of parental male Korean Oge (mKO) and female White Legrhon (fWL) were subjected to whole-genome sequencing. Hybrid single embryos between mKO and fWL at the oocyte, zygote, and EGK.X stages from each parent were subjected to whole-transcriptome sequencing. Allelic expression in the hybrid embryos was examined based on breed-specific SNPs. (B) A multidimensional scaling (MDS) plot based on log2 TMM normalized gene expressions of the whole transcriptome in pre-oviposited chicken embryos. Biological triplicates of single embryo were clustered and three developmental stages were clearly distinct. (C) The number of significantly detected long transcripts (e.g., mRNAs and lincRNAs)

detected by comparing gene expression among single oocyte, zygote, and EGK.X embryo (FDR-adjusted $P < 0.05$).

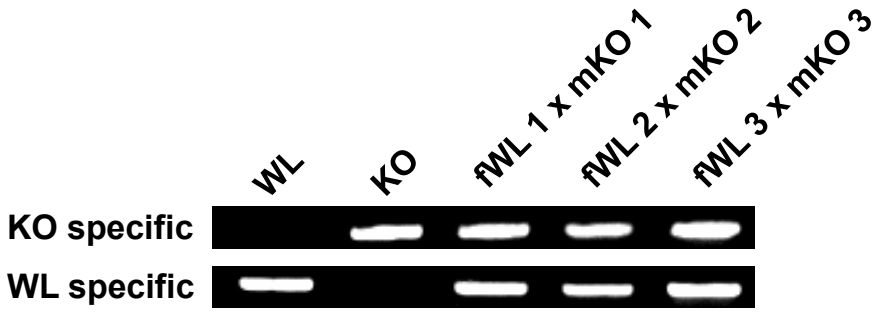


**Figure 6-7. The confirmation of hybrid embryo (Hamburger and Hamilton stage 4) between female White Legrhon (fWL) and male Korean Oge (mKO) using breed-specific primers.** WL, White Leghorn control; KO, Korean Oge control.

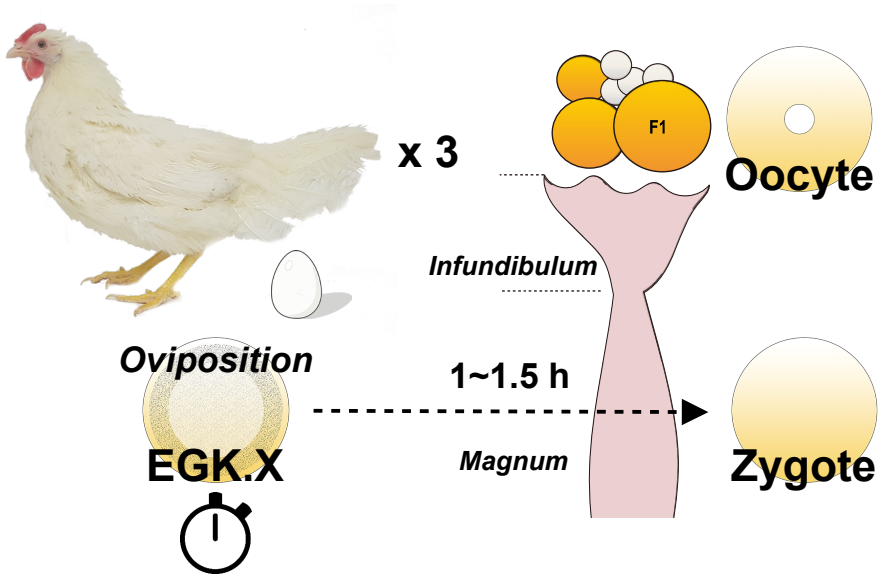


**Figure 6-8. A schematic diagram of single oocyte, zygote and EGK.X embryo acquisition from one hen at the same day.** At the day of acquiring embryos, EGK.X blastoderm was acquired at oviposition and the time was checked (0 h). Approximately 1 ~ 1.5 h after oviposition, pre-ovulatory F1 oocyte in the ovary and zygote in the magnum were simultaneously collected.

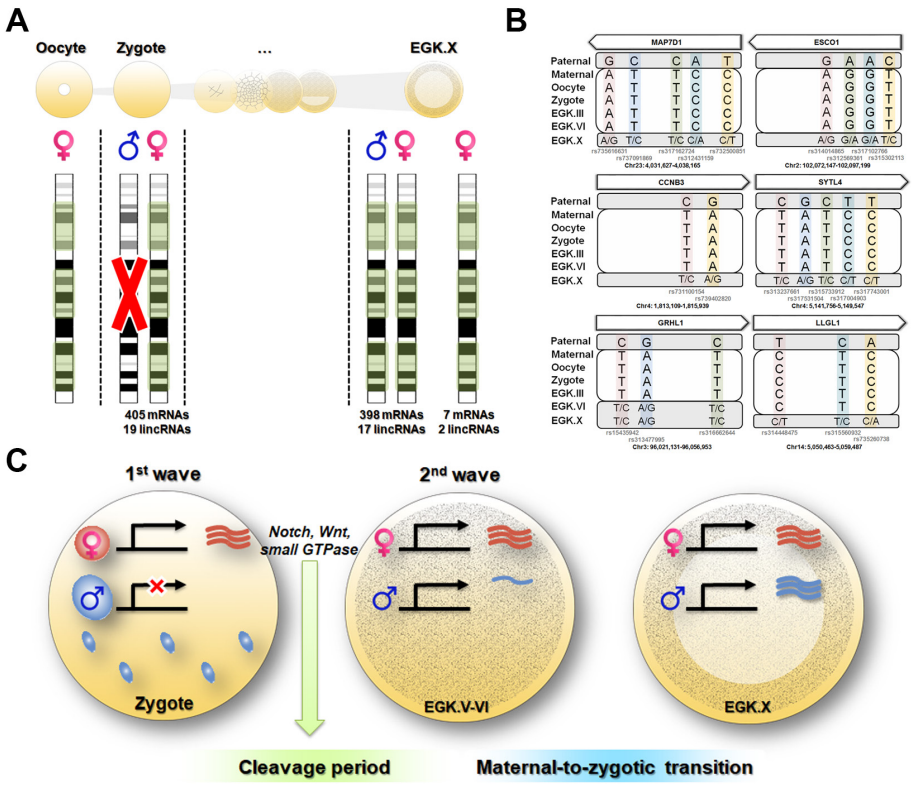


**Figure 6-9. Maternal genome activation (MGA) by 1st wave of transcriptional activation in chicken zygote.** (A) Determination of parental allelic expression from the zygote stage. Only maternal alleles were observed in transcripts induced by 1st activation. These maternally derived up-regulated genes showed biallelic expression after EGK.X. (B) Validation of 1st transcription-induced maternal allelic expression by Sanger sequencing. The maternal transcription profile after the 1st wave changed to biallelic expression between EGK.VI and EGK.X after the 2nd activation. (C) Schematic summary of genome activation during chicken early development. Only MGA occurred after fertilization and may regulate the cleavage period.

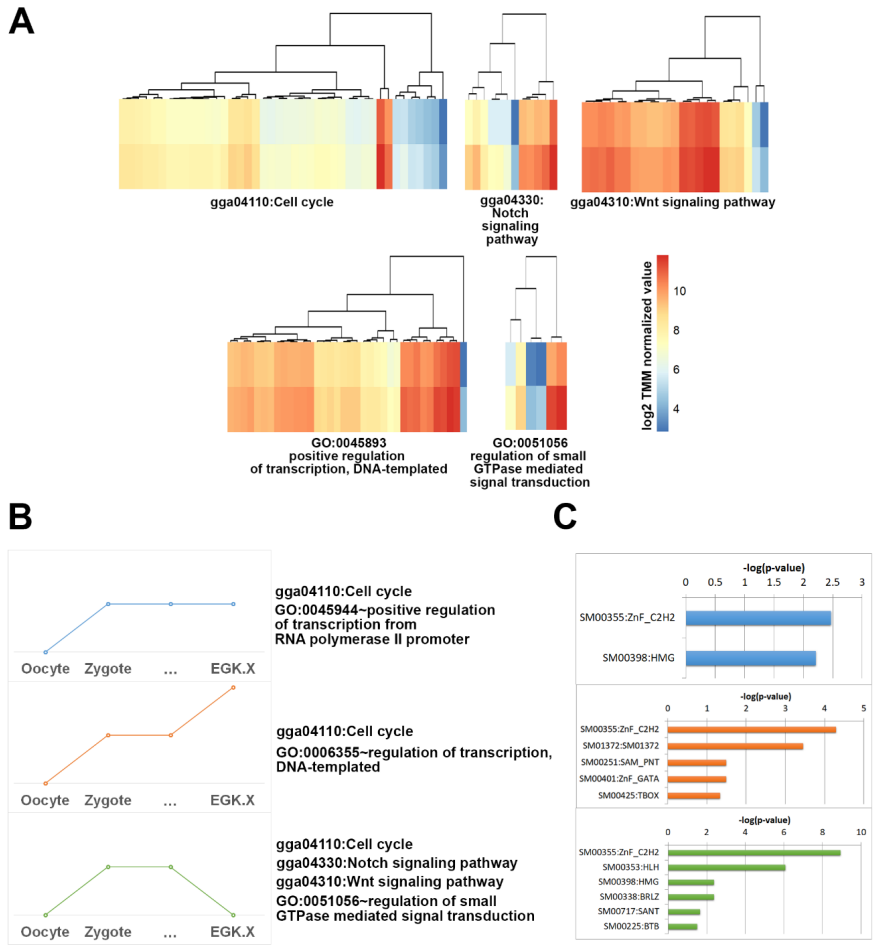


**Figure 6-10. Functional classification of genes by maternal genome activation during 1st wave of transcriptional activation and tracing through early development.** (A) The heatmaps showing expression patterns of the significantly up-regulated transcripts between oocyte and zygote (FDR-adjusted $P < 0.05$ and logFC > 0) in terms of biological processes in GO and KEGG pathway enrichment. (B) The dynamics of maternal derived up-regulated transcripts by 1st wave and biological terms during MZT. (C) SMART domain analysis of transcription factors up-regulated by 1st wave.

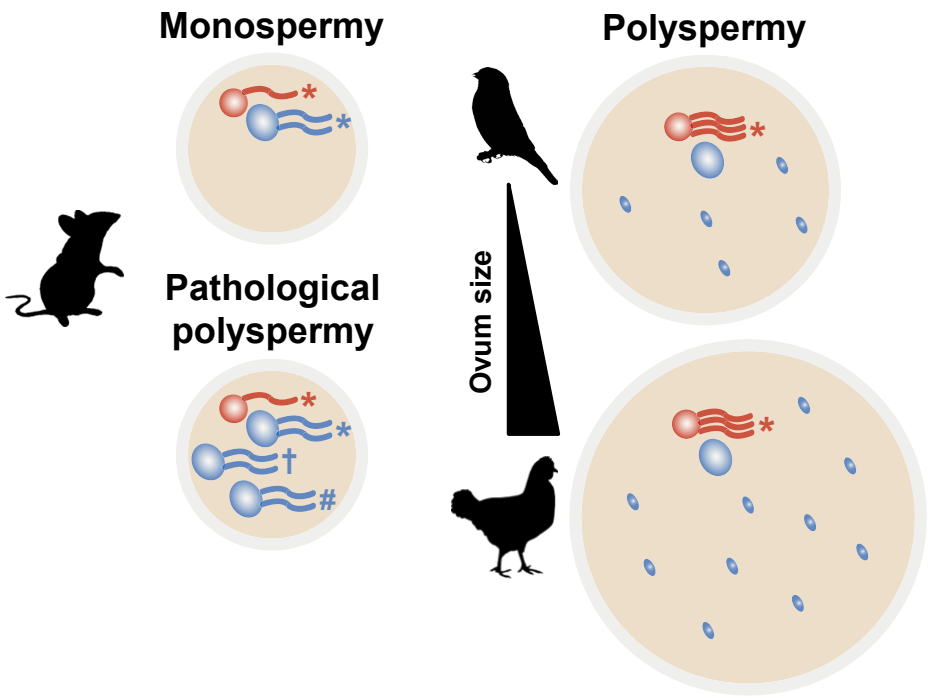


**Figure 6-11. The hypothetical diagram for avian polyspermy and the only maternal genome activation after fertilization.** Maternal genome activation in avian zygote would be needed for overcoming excessive genetic materials and genetic instability due to variability of polyspermy compared to mammal. *, †, and # indicate different type of transcripts by multiple sperm. Large red circle, maternal pronucleus (PN); large blue circle, paternal PN; small blue circle, supernumerary sperm nucleus in avian species.

**Table 6-1. The gene list and expressions of transcripts for exon-intron PCR**

**The genes of 1st wave of activation**

| Gene symbol | Associated Gene Name | FDR_Oocyte_Zygote_Intron | logFC_Oocyte_Zygote_Intron |
|---|---|---|---|
| ENSGALG00000023195 | **DLX6** | 4.16E-14 | 6.963060938 |
| ENSGALG00000005909 | **GATA2** | 1.62E-16 | 3.952010232 |
| ENSGALG00000006810 | **ZIC4** | 1.26E-09 | 4.371072578 |

**The genes of 2nd wave of activation**

| Gene symbol | Associated Gene Name | FDR_EGK.III_EGK.VI_Intron | logFC_EGK.III_EGK.VI_Intron |
|---|---|---|---|
| ENSGALG00000000839 | **WNT11** | 4.68E-27 | 11.27729466 |
| ENSGALG00000005396 | **WNT3A** | 3.83E-28 | 10.56970018 |
| ENSGALG00000024487 | **C8ORF22** | 5.27E-16 | 9.930131588 |

**Table 6-2. The total RNA quantity of single chicken early embryo**

| No. | Sample ID. | Ribogreen | Volume | Amount |
|---|---|---|---|---|
| | | Con. (ng/ul) | (ul) | (ng) |
| 1 | fWL1 Oocyte | 98.5 | 20 | 1970 |
| 2 | fWL2 Oocyte | 88.6 | 20 | 1772 |
| 3 | fWL3 Oocyte | 136.3 | 20 | 2726 |
| 4 | fWL 1 x mKO 1 Zygote | 77.8 | 20 | 1556 |
| 5 | fWL 2 x mKO 2 Zygote | 116.9 | 20 | 2338 |
| 6 | fWL 3 x mKO 3 Zygote | 93.1 | 20 | 1862 |
| 7 | fWL 1 x mKO 1 EGK.X | 114.9 | 20 | 2298 |
| 8 | fWL 2 x mKO 2 EGK.X | 121.4 | 20 | 2428 |
| 9 | fWL 3 x mKO 3 EGK.X | 100.4 | 20 | 2008 |

**Table 6-3. The up-regulated intronic expression between single oocyte and zygote (FDR adjusted P < 0.05)**

| | Oocyte → Zygote | Zygote → EGK.X |
|---|---|---|
| Up | 5312 | 5438 |
| Down | 3328 | 3843 |

**Table 6-4. The variant calling of single hybrid embryo RNA-seq for determining of which parental allele expressed**

| Gene symbol | SNP Counts | Cluster |
|---|---|---|
| ENSGALG00000000109 | 2 | Maternal_Bi |
| ENSGALG00000000129 | 3 | Maternal_Bi |
| ENSGALG00000000177 | 4 | Maternal_Bi |
| ENSGALG00000000234 | 3 | Maternal_Bi |
| ENSGALG00000000258 | 5 | Maternal_Bi |
| ENSGALG00000000299 | 4 | Maternal_Bi |
| ENSGALG00000000305 | 3 | Maternal_Bi |
| ENSGALG00000000513 | 4 | Maternal_Bi |
| ENSGALG00000000516 | 17 | Maternal_Bi |
| ENSGALG00000000540 | 3 | Maternal_Bi |
| ENSGALG00000000581 | 3 | Maternal_Bi |
| ENSGALG00000000608 | 7 | Maternal_Bi |
| ENSGALG00000000632 | 5 | Maternal_Bi |
| ENSGALG00000000638 | 3 | Maternal_Bi |
| ENSGALG00000000657 | 5 | Maternal_Bi |
| ENSGALG00000000671 | 5 | Maternal_Bi |
| ENSGALG00000000693 | 2 | Maternal_Bi |
| ENSGALG00000000699 | 6 | Maternal_Bi |
| ENSGALG00000000701 | 3 | Maternal_Bi |
| ENSGALG00000000713 | 2 | Maternal_Bi |
| ENSGALG00000000721 | 7 | Maternal_Bi |
| ENSGALG00000000771 | 2 | Maternal_Bi |
| ENSGALG00000000794 | 4 | Maternal_Bi |
| ENSGALG00000000816 | 2 | Maternal_Bi |
| ENSGALG00000000939 | 2 | Maternal_Bi |
| ENSGALG00000000982 | 2 | Maternal_Bi |
| ENSGALG00000001020 | 5 | Maternal_Bi |
| ENSGALG00000001126 | 5 | Maternal_Bi |
| ENSGALG00000001182 | 2 | Maternal_Bi |
| ENSGALG00000001231 | 4 | Maternal_Bi |
| ENSGALG00000001270 | 2 | Maternal_Bi |
| ENSGALG00000001420 | 2 | Maternal_Bi |
| ENSGALG00000001479 | 2 | Maternal_Bi |
| ENSGALG00000001482 | 4 | Maternal_Bi |
| ENSGALG00000001492 | 4 | Maternal_Bi |
| ENSGALG00000001516 | 3 | Maternal_Bi |
| ENSGALG00000001522 | 2 | Maternal_Bi |
| ENSGALG00000001536 | 8 | Maternal_Bi |
| ENSGALG00000001540 | 4 | Maternal_Bi |
| ENSGALG00000001672 | 6 | Maternal_Bi |
| ENSGALG00000001692 | 3 | Maternal_Bi |
| ENSGALG00000001722 | 2 | Maternal_Bi |
| ENSGALG00000001732 | 2 | Maternal_Bi |
| ENSGALG00000001775 | 8 | Maternal_Bi |
| ENSGALG00000001779 | 4 | Maternal_Bi |
| ENSGALG00000001907 | 2 | Maternal_Bi |

| | | |
|---|---|---|
| ENSGALG00000001910 | 2 | Maternal_Bi |
| ENSGALG00000002030 | 2 | Maternal_Bi |
| ENSGALG00000002041 | 7 | Maternal_Bi |
| ENSGALG00000002144 | 6 | Maternal_Bi |
| ENSGALG00000002158 | 2 | Maternal_Bi |
| ENSGALG00000002242 | 3 | Maternal_Bi |
| ENSGALG00000002272 | 2 | Maternal_Bi |
| ENSGALG00000002284 | 8 | Maternal_Bi |
| ENSGALG00000002336 | 2 | Maternal_Bi |
| ENSGALG00000002354 | 2 | Maternal_Bi |
| ENSGALG00000002398 | 4 | Maternal_Bi |
| ENSGALG00000002412 | 2 | Maternal_Bi |
| ENSGALG00000002447 | 2 | Maternal_Bi |
| ENSGALG00000002461 | 2 | Maternal_Bi |
| ENSGALG00000002489 | 2 | Maternal_Bi |
| ENSGALG00000002527 | 2 | Maternal_Bi |
| ENSGALG00000002553 | 4 | Maternal_Bi |
| ENSGALG00000002591 | 2 | Maternal_Bi |
| ENSGALG00000002622 | 2 | Maternal_Bi |
| ENSGALG00000002695 | 2 | Maternal_Bi |
| ENSGALG00000002699 | 3 | Maternal_Bi |
| ENSGALG00000002712 | 3 | Maternal_Bi |
| ENSGALG00000002781 | 3 | Maternal_Bi |
| ENSGALG00000002818 | 2 | Maternal_Bi |
| ENSGALG00000002839 | 3 | Maternal_Bi |
| ENSGALG00000002880 | 2 | Maternal_Bi |
| ENSGALG00000002932 | 2 | Maternal_Bi |
| ENSGALG00000002943 | 7 | Maternal_Bi |
| ENSGALG00000002971 | 8 | Maternal_Bi |
| ENSGALG00000002984 | 3 | Maternal_Bi |
| ENSGALG00000002985 | 8 | Maternal_Bi |
| ENSGALG00000003011 | 2 | Maternal_Bi |
| ENSGALG00000003168 | 3 | Maternal_Bi |
| ENSGALG00000003194 | 3 | Maternal_Bi |
| ENSGALG00000003392 | 3 | Maternal_Bi |
| ENSGALG00000003527 | 2 | Maternal_Bi |
| ENSGALG00000003620 | 4 | Maternal_Bi |
| ENSGALG00000003663 | 11 | Maternal_Bi |
| ENSGALG00000003712 | 2 | Maternal_Bi |
| ENSGALG00000003719 | 4 | Maternal_Bi |
| ENSGALG00000003739 | 4 | Maternal_Bi |
| ENSGALG00000003959 | 4 | Maternal_Bi |
| ENSGALG00000004291 | 2 | Maternal_Bi |
| ENSGALG00000004378 | 6 | Maternal_Bi |
| ENSGALG00000004471 | 5 | Maternal_Bi |
| ENSGALG00000004575 | 3 | Maternal_Bi |
| ENSGALG00000004578 | 2 | Maternal_Bi |
| ENSGALG00000004584 | 19 | Maternal_Bi |
| ENSGALG00000004599 | 5 | Maternal_Bi |
| ENSGALG00000004602 | 3 | Maternal |
| ENSGALG00000004612 | 8 | Maternal_Bi |
| ENSGALG00000004637 | 3 | Maternal_Bi |
| ENSGALG00000004661 | 4 | Maternal_Bi |

| | | |
|---|---|---|
| ENSGALG00000004670 | 3 | Maternal_Bi |
| ENSGALG00000004679 | 3 | Maternal_Bi |
| ENSGALG00000004691 | 6 | Maternal_Bi |
| ENSGALG00000004728 | 3 | Maternal_Bi |
| ENSGALG00000004750 | 2 | Maternal_Bi |
| ENSGALG00000004800 | 4 | Maternal_Bi |
| ENSGALG00000004886 | 2 | Maternal_Bi |
| ENSGALG00000004958 | 2 | Maternal_Bi |
| ENSGALG00000004959 | 2 | Maternal_Bi |
| ENSGALG00000004960 | 4 | Maternal_Bi |
| ENSGALG00000005046 | 2 | Maternal_Bi |
| ENSGALG00000005054 | 3 | Maternal_Bi |
| ENSGALG00000005071 | 3 | Maternal_Bi |
| ENSGALG00000005205 | 2 | Maternal_Bi |
| ENSGALG00000005269 | 2 | Maternal_Bi |
| ENSGALG00000005337 | 2 | Maternal_Bi |
| ENSGALG00000005367 | 2 | Maternal_Bi |
| ENSGALG00000005673 | 2 | Maternal_Bi |
| ENSGALG00000005682 | 4 | Maternal_Bi |
| ENSGALG00000005703 | 4 | Maternal_Bi |
| ENSGALG00000005846 | 2 | Maternal_Bi |
| ENSGALG00000005855 | 7 | Maternal_Bi |
| ENSGALG00000005904 | 2 | Maternal_Bi |
| ENSGALG00000006016 | 5 | Maternal_Bi |
| ENSGALG00000006047 | 2 | Maternal_Bi |
| ENSGALG00000006051 | 2 | Maternal_Bi |
| ENSGALG00000006080 | 3 | Maternal_Bi |
| ENSGALG00000006118 | 2 | Maternal_Bi |
| ENSGALG00000006122 | 2 | Maternal_Bi |
| ENSGALG00000006281 | 3 | Maternal_Bi |
| ENSGALG00000006298 | 3 | Maternal_Bi |
| ENSGALG00000006305 | 2 | Maternal_Bi |
| ENSGALG00000006401 | 2 | Maternal_Bi |
| ENSGALG00000006405 | 2 | Maternal_Bi |
| ENSGALG00000006418 | 4 | Maternal_Bi |
| ENSGALG00000006437 | 8 | Maternal_Bi |
| ENSGALG00000006472 | 2 | Maternal_Bi |
| ENSGALG00000006561 | 2 | Maternal_Bi |
| ENSGALG00000006630 | 3 | Maternal_Bi |
| ENSGALG00000006646 | 4 | Maternal_Bi |
| ENSGALG00000006658 | 2 | Maternal_Bi |
| ENSGALG00000006694 | 2 | Maternal_Bi |
| ENSGALG00000006773 | 2 | Maternal_Bi |
| ENSGALG00000006775 | 3 | Maternal_Bi |
| ENSGALG00000006786 | 5 | Maternal_Bi |
| ENSGALG00000006788 | 2 | Maternal_Bi |
| ENSGALG00000006791 | 2 | Maternal_Bi |
| ENSGALG00000006925 | 3 | Maternal_Bi |
| ENSGALG00000006960 | 2 | Maternal_Bi |
| ENSGALG00000007089 | 3 | Maternal_Bi |
| ENSGALG00000007119 | 5 | Maternal_Bi |
| ENSGALG00000007128 | 2 | Maternal_Bi |
| ENSGALG00000007151 | 4 | Maternal_Bi |

| | | |
|---|---|---|
| ENSGALG00000007276 | 5 | Maternal_Bi |
| ENSGALG00000007307 | 2 | Maternal_Bi |
| ENSGALG00000007334 | 2 | Maternal_Bi |
| ENSGALG00000007354 | 2 | Maternal_Bi |
| ENSGALG00000007373 | 8 | Maternal_Bi |
| ENSGALG00000007412 | 6 | Maternal_Bi |
| ENSGALG00000007439 | 4 | Maternal_Bi |
| ENSGALG00000007447 | 4 | Maternal_Bi |
| ENSGALG00000007462 | 2 | Maternal_Bi |
| ENSGALG00000007591 | 2 | Maternal_Bi |
| ENSGALG00000007644 | 2 | Maternal_Bi |
| ENSGALG00000007665 | 4 | Maternal_Bi |
| ENSGALG00000007686 | 4 | Maternal_Bi |
| ENSGALG00000007829 | 2 | Maternal_Bi |
| ENSGALG00000007928 | 2 | Maternal_Bi |
| ENSGALG00000008103 | 2 | Maternal_Bi |
| ENSGALG00000008162 | 3 | Maternal_Bi |
| ENSGALG00000008174 | 3 | Maternal_Bi |
| ENSGALG00000008180 | 2 | Maternal_Bi |
| ENSGALG00000008217 | 2 | Maternal_Bi |
| ENSGALG00000008219 | 3 | Maternal_Bi |
| ENSGALG00000008229 | 2 | Maternal_Bi |
| ENSGALG00000008254 | 3 | Maternal_Bi |
| ENSGALG00000008257 | 7 | Maternal_Bi |
| ENSGALG00000008410 | 8 | Maternal_Bi |
| ENSGALG00000008561 | 4 | Maternal_Bi |
| ENSGALG00000008618 | 2 | Maternal_Bi |
| ENSGALG00000008764 | 2 | Maternal_Bi |
| ENSGALG00000008771 | 2 | Maternal_Bi |
| ENSGALG00000008803 | 3 | Maternal_Bi |
| ENSGALG00000008806 | 3 | Maternal_Bi |
| ENSGALG00000008813 | 2 | Maternal_Bi |
| ENSGALG00000008819 | 3 | Maternal_Bi |
| ENSGALG00000008838 | 3 | Maternal_Bi |
| ENSGALG00000008933 | 2 | Maternal_Bi |
| ENSGALG00000009001 | 4 | Maternal_Bi |
| ENSGALG00000009031 | 2 | Maternal_Bi |
| ENSGALG00000009042 | 2 | Maternal_Bi |
| ENSGALG00000009057 | 2 | Maternal_Bi |
| ENSGALG00000009083 | 9 | Maternal_Bi |
| ENSGALG00000009158 | 3 | Maternal_Bi |
| ENSGALG00000009252 | 5 | Maternal_Bi |
| ENSGALG00000009283 | 3 | Maternal_Bi |
| ENSGALG00000009381 | 5 | Maternal_Bi |
| ENSGALG00000009437 | 2 | Maternal_Bi |
| ENSGALG00000009452 | 2 | Maternal_Bi |
| ENSGALG00000009533 | 2 | Maternal_Bi |
| ENSGALG00000009626 | 2 | Maternal_Bi |
| ENSGALG00000009837 | 2 | Maternal_Bi |
| ENSGALG00000009840 | 2 | Maternal_Bi |
| ENSGALG00000009854 | 3 | Maternal_Bi |
| ENSGALG00000010001 | 3 | Maternal_Bi |
| ENSGALG00000010020 | 2 | Maternal_Bi |

| | | |
|---|---|---|
| ENSGALG00000010030 | 2 | Maternal_Bi |
| ENSGALG00000010127 | 6 | Maternal_Bi |
| ENSGALG00000010164 | 2 | Maternal_Bi |
| ENSGALG00000010166 | 3 | Maternal_Bi |
| ENSGALG00000010234 | 5 | Maternal_Bi |
| ENSGALG00000010235 | 2 | Maternal_Bi |
| ENSGALG00000010255 | 2 | Maternal_Bi |
| ENSGALG00000010265 | 12 | Maternal_Bi |
| ENSGALG00000010341 | 6 | Maternal_Bi |
| ENSGALG00000010360 | 6 | Maternal_Bi |
| ENSGALG00000010448 | 4 | Maternal_Bi |
| ENSGALG00000010541 | 3 | Maternal_Bi |
| ENSGALG00000010677 | 3 | Maternal_Bi |
| ENSGALG00000010719 | 7 | Maternal_Bi |
| ENSGALG00000010738 | 4 | Maternal_Bi |
| ENSGALG00000010766 | 3 | Maternal_Bi |
| ENSGALG00000010827 | 3 | Maternal_Bi |
| ENSGALG00000010934 | 2 | Maternal_Bi |
| ENSGALG00000010940 | 3 | Maternal |
| ENSGALG00000011015 | 2 | Maternal_Bi |
| ENSGALG00000011068 | 9 | Maternal_Bi |
| ENSGALG00000011481 | 6 | Maternal_Bi |
| ENSGALG00000011488 | 4 | Maternal_Bi |
| ENSGALG00000011517 | 2 | Maternal_Bi |
| ENSGALG00000011639 | 3 | Maternal_Bi |
| ENSGALG00000011643 | 9 | Maternal_Bi |
| ENSGALG00000011655 | 2 | Maternal_Bi |
| ENSGALG00000011689 | 4 | Maternal_Bi |
| ENSGALG00000011811 | 3 | Maternal_Bi |
| ENSGALG00000011833 | 3 | Maternal_Bi |
| ENSGALG00000011893 | 3 | Maternal_Bi |
| ENSGALG00000012025 | 2 | Maternal_Bi |
| ENSGALG00000012044 | 4 | Maternal_Bi |
| ENSGALG00000012048 | 3 | Maternal_Bi |
| ENSGALG00000012089 | 6 | Maternal_Bi |
| ENSGALG00000012171 | 2 | Maternal_Bi |
| ENSGALG00000012244 | 2 | Maternal_Bi |
| ENSGALG00000012256 | 9 | Maternal_Bi |
| ENSGALG00000012263 | 5 | Maternal_Bi |
| ENSGALG00000012321 | 3 | Maternal_Bi |
| ENSGALG00000012395 | 3 | Maternal_Bi |
| ENSGALG00000012455 | 3 | Maternal_Bi |
| ENSGALG00000012588 | 3 | Maternal_Bi |
| ENSGALG00000012615 | 2 | Maternal_Bi |
| ENSGALG00000012680 | 3 | Maternal_Bi |
| ENSGALG00000012898 | 2 | Maternal_Bi |
| ENSGALG00000012903 | 3 | Maternal_Bi |
| ENSGALG00000012909 | 4 | Maternal_Bi |
| ENSGALG00000012988 | 2 | Maternal_Bi |
| ENSGALG00000013005 | 3 | Maternal_Bi |
| ENSGALG00000013053 | 2 | Maternal_Bi |
| ENSGALG00000013063 | 2 | Maternal_Bi |
| ENSGALG00000013117 | 4 | Maternal |

| | | |
|---|---|---|
| ENSGALG00000013493 | 2 | Maternal_Bi |
| ENSGALG00000013573 | 12 | Maternal_Bi |
| ENSGALG00000013605 | 2 | Maternal_Bi |
| ENSGALG00000013647 | 2 | Maternal_Bi |
| ENSGALG00000013677 | 6 | Maternal_Bi |
| ENSGALG00000013773 | 5 | Maternal_Bi |
| ENSGALG00000013795 | 3 | Maternal_Bi |
| ENSGALG00000013803 | 2 | Maternal_Bi |
| ENSGALG00000013804 | 2 | Maternal_Bi |
| ENSGALG00000014036 | 3 | Maternal_Bi |
| ENSGALG00000014163 | 3 | Maternal_Bi |
| ENSGALG00000014270 | 4 | Maternal_Bi |
| ENSGALG00000014281 | 18 | Maternal_Bi |
| ENSGALG00000014337 | 2 | Maternal_Bi |
| ENSGALG00000014404 | 5 | Maternal_Bi |
| ENSGALG00000014408 | 3 | Maternal_Bi |
| ENSGALG00000014649 | 2 | Maternal_Bi |
| ENSGALG00000014652 | 2 | Maternal_Bi |
| ENSGALG00000014684 | 3 | Maternal_Bi |
| ENSGALG00000014779 | 2 | Maternal |
| ENSGALG00000014784 | 15 | Maternal_Bi |
| ENSGALG00000014788 | 4 | Maternal_Bi |
| ENSGALG00000014815 | 2 | Maternal_Bi |
| ENSGALG00000014836 | 4 | Maternal_Bi |
| ENSGALG00000014852 | 2 | Maternal_Bi |
| ENSGALG00000014865 | 3 | Maternal_Bi |
| ENSGALG00000014923 | 2 | Maternal |
| ENSGALG00000014936 | 4 | Maternal_Bi |
| ENSGALG00000014960 | 3 | Maternal_Bi |
| ENSGALG00000014974 | 7 | Maternal_Bi |
| ENSGALG00000014982 | 4 | Maternal_Bi |
| ENSGALG00000015004 | 17 | Maternal_Bi |
| ENSGALG00000015011 | 6 | Maternal_Bi |
| ENSGALG00000015027 | 4 | Maternal_Bi |
| ENSGALG00000015030 | 11 | Maternal_Bi |
| ENSGALG00000015038 | 3 | Maternal_Bi |
| ENSGALG00000015059 | 3 | Maternal_Bi |
| ENSGALG00000015064 | 3 | Maternal_Bi |
| ENSGALG00000015069 | 2 | Maternal_Bi |
| ENSGALG00000015071 | 6 | Maternal_Bi |
| ENSGALG00000015086 | 2 | Maternal_Bi |
| ENSGALG00000015093 | 3 | Maternal_Bi |
| ENSGALG00000015105 | 2 | Maternal_Bi |
| ENSGALG00000015106 | 5 | Maternal |
| ENSGALG00000015152 | 2 | Maternal_Bi |
| ENSGALG00000015170 | 4 | Maternal_Bi |
| ENSGALG00000015179 | 4 | Maternal_Bi |
| ENSGALG00000015331 | 5 | Maternal_Bi |
| ENSGALG00000015406 | 2 | Maternal |
| ENSGALG00000015471 | 2 | Maternal_Bi |
| ENSGALG00000015479 | 2 | Maternal_Bi |
| ENSGALG00000015486 | 4 | Maternal_Bi |
| ENSGALG00000015524 | 2 | Maternal_Bi |

| | | |
|---|---|---|
| ENSGALG00000015539 | 8 | Maternal_Bi |
| ENSGALG00000015556 | 2 | Maternal_Bi |
| ENSGALG00000015589 | 3 | Maternal_Bi |
| ENSGALG00000015611 | 3 | Maternal_Bi |
| ENSGALG00000015665 | 3 | Maternal_Bi |
| ENSGALG00000015836 | 2 | Maternal_Bi |
| ENSGALG00000015916 | 2 | Maternal_Bi |
| ENSGALG00000015955 | 2 | Maternal_Bi |
| ENSGALG00000015957 | 2 | Maternal_Bi |
| ENSGALG00000015980 | 5 | Maternal_Bi |
| ENSGALG00000015984 | 2 | Maternal_Bi |
| ENSGALG00000015985 | 2 | Maternal_Bi |
| ENSGALG00000016026 | 4 | Maternal_Bi |
| ENSGALG00000016034 | 2 | Maternal_Bi |
| ENSGALG00000016070 | 7 | Maternal_Bi |
| ENSGALG00000016092 | 5 | Maternal_Bi |
| ENSGALG00000016125 | 3 | Maternal_Bi |
| ENSGALG00000016174 | 4 | Maternal_Bi |
| ENSGALG00000016282 | 7 | Maternal_Bi |
| ENSGALG00000016289 | 4 | Maternal_Bi |
| ENSGALG00000016320 | 4 | Maternal_Bi |
| ENSGALG00000016409 | 2 | Maternal_Bi |
| ENSGALG00000016424 | 4 | Maternal_Bi |
| ENSGALG00000016427 | 3 | Maternal_Bi |
| ENSGALG00000016438 | 2 | Maternal_Bi |
| ENSGALG00000016439 | 3 | Maternal_Bi |
| ENSGALG00000016456 | 2 | Maternal_Bi |
| ENSGALG00000016576 | 5 | Maternal_Bi |
| ENSGALG00000016605 | 2 | Maternal_Bi |
| ENSGALG00000016608 | 3 | Maternal_Bi |
| ENSGALG00000016609 | 10 | Maternal_Bi |
| ENSGALG00000016649 | 4 | Maternal_Bi |
| ENSGALG00000016686 | 4 | Maternal_Bi |
| ENSGALG00000016691 | 2 | Maternal_Bi |
| ENSGALG00000016702 | 4 | Maternal_Bi |
| ENSGALG00000016719 | 2 | Maternal_Bi |
| ENSGALG00000016732 | 3 | Maternal_Bi |
| ENSGALG00000016860 | 2 | Maternal_Bi |
| ENSGALG00000016870 | 5 | Maternal_Bi |
| ENSGALG00000016991 | 2 | Maternal_Bi |
| ENSGALG00000017071 | 2 | Maternal_Bi |
| ENSGALG00000017112 | 2 | Maternal_Bi |
| ENSGALG00000017128 | 7 | Maternal_Bi |
| ENSGALG00000017159 | 4 | Maternal_Bi |
| ENSGALG00000017392 | 4 | Maternal_Bi |
| ENSGALG00000018800 | 3 | Maternal_Bi |
| ENSGALG00000018967 | 8 | Maternal_Bi |
| ENSGALG00000019201 | 2 | Maternal_Bi |
| ENSGALG00000019211 | 4 | Maternal_Bi |
| ENSGALG00000019228 | 2 | Maternal_Bi |
| ENSGALG00000019336 | 4 | Maternal_Bi |
| ENSGALG00000019710 | 3 | Maternal_Bi |
| ENSGALG00000019891 | 4 | Maternal_Bi |

| | | |
|---|---|---|
| ENSGALG00000020003 | 2 | Maternal_Bi |
| ENSGALG00000020561 | 7 | Maternal_Bi |
| ENSGALG00000020729 | 2 | Maternal_Bi |
| ENSGALG00000020878 | 3 | Maternal_Bi |
| ENSGALG00000020893 | 16 | Maternal_Bi |
| ENSGALG00000021511 | 4 | Maternal_Bi |
| ENSGALG00000021634 | 5 | Maternal_Bi |
| ENSGALG00000022490 | 2 | Maternal_Bi |
| ENSGALG00000022891 | 2 | Maternal_Bi |
| ENSGALG00000022920 | 5 | Maternal_Bi |
| ENSGALG00000023616 | 2 | Maternal_Bi |
| ENSGALG00000023683 | 2 | Maternal_Bi |
| ENSGALG00000023731 | 3 | Maternal_Bi |
| ENSGALG00000025712 | 2 | Maternal_Bi |
| ENSGALG00000025810 | 2 | Maternal_Bi |
| ENSGALG00000026113 | 3 | Maternal_Bi |
| ENSGALG00000026144 | 3 | Maternal_Bi |
| ENSGALG00000026159 | 2 | Maternal_Bi |
| ENSGALG00000026231 | 2 | Maternal_Bi |
| ENSGALG00000026291 | 9 | Maternal_Bi |
| ENSGALG00000026460 | 2 | Maternal_Bi |
| ENSGALG00000026505 | 2 | Maternal_Bi |
| ENSGALG00000026542 | 4 | Maternal_Bi |
| ENSGALG00000026757 | 3 | Maternal_Bi |
| ENSGALG00000026900 | 3 | Maternal_Bi |
| ENSGALG00000027143 | 3 | Maternal_Bi |
| ENSGALG00000027190 | 11 | Maternal_Bi |
| ENSGALG00000027191 | 2 | Maternal_Bi |
| ENSGALG00000027482 | 2 | Maternal_Bi |
| ENSGALG00000027871 | 8 | Maternal_Bi |
| ENSGALG00000027874 | 2 | Maternal_Bi |
| ENSGALG00000028015 | 2 | Maternal_Bi |
| ENSGALG00000028084 | 4 | Maternal_Bi |
| ENSGALG00000028148 | 10 | Maternal_Bi |
| ENSGALG00000028440 | 2 | Maternal_Bi |
| ENSGALG00000028543 | 4 | Maternal_Bi |
| ENSGALG00000028691 | 3 | Maternal_Bi |
| ENSGALG00000028696 | 2 | Maternal_Bi |
| ENSGALG00000028821 | 2 | Maternal_Bi |
| ENSGALG00000028967 | 7 | Maternal_Bi |
| ENSGALG00000029109 | 2 | Maternal_Bi |
| ALDBGALG0000000155 | 3 | Maternal_Bi |
| ALDBGALG0000000159 | 4 | Maternal_Bi |
| ALDBGALG0000000492 | 2 | Maternal_Bi |
| ALDBGALG0000001196 | 3 | Maternal_Bi |
| ALDBGALG0000001304 | 2 | Maternal_Bi |
| ALDBGALG0000001553 | 2 | Maternal_Bi |
| ALDBGALG0000001582 | 3 | Maternal_Bi |
| ALDBGALG0000001742 | 5 | Maternal_Bi |
| ALDBGALG0000001820 | 9 | Maternal |
| ALDBGALG0000002562 | 2 | Maternal_Bi |
| ALDBGALG0000002716 | 3 | Maternal_Bi |
| ALDBGALG0000003224 | 3 | Maternal_Bi |

| | | |
|---|---|---|
| ALDBGALG0000003328 | 2 | Maternal_Bi |
| ALDBGALG0000003669 | 4 | Maternal_Bi |
| ALDBGALG0000003866 | 2 | Maternal_Bi |
| ALDBGALG0000004579 | 3 | Maternal_Bi |
| ALDBGALG0000005157 | 2 | Maternal |
| ALDBGALG0000005226 | 3 | Maternal_Bi |
| ALDBGALG0000005723 | 2 | Maternal_Bi |

**Table 6-5. The gene list and expressions of genotyped transcripts by Sanger sequencing**

| Gene symbol | Associated Gene Name | FDR_Oocyte_Zygote_DEG | logFC_Oocyte_Zygote_DEG | FDR_Oocyte_Zygote_Intron | logFC_Oocyte_Zygote_Intron |
|---|---|---|---|---|---|
| ENSGALG00000021634 | **MAP7D1** | 1.08E-05 | 1.023546155 | 2.12E-23 | 2.63019146 |
| ENSGALG00000025810 | **CCNB3** | 2.39E-08 | 0.816494756 | 2.03E-13 | 2.372385693 |
| ENSGALG00000016439 | **GRHL1** | 3.35E-02 | 0.708823487 | 1.58E-15 | 2.617955026 |
| ENSGALG00000014936 | **ESCO1** | 1.38E-03 | 0.700238021 | 4.14E-11 | 1.550483786 |
| ENSGALG00000006786 | **SYTL4** | 1.39E-02 | 0.653131702 | 4.15E-14 | 2.249571462 |
| ENSGALG00000014408 | **LLGL1** | 1.38E-02 | 0.650003174 | 4.73E-14 | 2.300090484 |

**Table 6-6. Significantly detected biological processes of GO and KEGG pathways based on up-regulated DEGs between single oocyte and zygote**

**Significantly detected BP terms of GO based on up-regulated DEGs between oocyte and zygote (enrichment test P-value < 0.05)**

| GO term | # of genes | % | P-value |
|---|---|---|---|
| DNA replication initiation | 11 | 0.790797987 | 3.11E-07 |
| DNA replication | 20 | 1.437814522 | 4.26E-06 |
| cellular response to DNA damage stimulus | 21 | 1.509705248 | 7.15E-05 |
| cell division | 21 | 1.509705248 | 7.15E-05 |
| DNA repair | 23 | 1.6534867 | 2.56E-04 |
| positive regulation of transcription from RNA polymerase II promoter | 68 | 4.888569375 | 3.62E-04 |
| chromosome segregation | 13 | 0.934579439 | 3.70E-04 |
| ATP-dependent chromatin remodeling | 8 | 0.575125809 | 4.38E-04 |
| cell cycle | 14 | 1.006470165 | 5.62E-04 |
| mitotic cytokinesis | 9 | 0.647016535 | 6.05E-04 |
| G1/S transition of mitotic cell cycle | 12 | 0.862688713 | 6.76E-04 |
| actin cytoskeleton organization | 16 | 1.150251618 | 7.95E-04 |
| regulation of transcription, DNA-templated | 61 | 4.385334292 | 1.74E-03 |
| peptidyl-serine phosphorylation | 20 | 1.437814522 | 1.84E-03 |
| negative regulation of transcription from RNA polymerase II promoter | 50 | 3.594536305 | 1.86E-03 |
| microtubule-based movement | 14 | 1.006470165 | 2.01E-03 |
| organ morphogenesis | 11 | 0.790797987 | 2.07E-03 |
| base-excision repair | 9 | 0.647016535 | 2.40E-03 |
| double-strand break repair via homologous recombination | 14 | 1.006470165 | 2.43E-03 |
| neural tube closure | 13 | 0.934579439 | 3.52E-03 |
| meiotic chromosome segregation | 4 | 0.287562904 | 4.37E-03 |
| double-strand break repair | 9 | 0.647016535 | 5.45E-03 |
| mismatch repair | 6 | 0.4313 | 5.53E-03 |

| | | | |
|---|---|---|---|
| | | 44357 | |
| somatic hypermutation of immunoglobulin genes | 5 | 0.3594 5363 | 6.17E-03 |
| mRNA catabolic process | 5 | 0.3594 5363 | 6.17E-03 |
| gastrulation with mouth forming second | 6 | 0.4313 44357 | 8.21E-03 |
| protein monoubiquitination | 7 | 0.5032 35083 | 8.43E-03 |
| mitotic chromosome condensation | 5 | 0.3594 5363 | 1.02E-02 |
| regulation of cell shape | 16 | 1.1502 51618 | 1.10E-02 |
| G2 DNA damage checkpoint | 6 | 0.4313 44357 | 1.17E-02 |
| regulation of small GTPase mediated signal transduction | 6 | 0.4313 44357 | 1.17E-02 |
| spindle assembly | 6 | 0.4313 44357 | 1.17E-02 |
| mRNA splicing, via spliceosome | 13 | 0.9345 79439 | 1.28E-02 |
| response to ionizing radiation | 9 | 0.6470 16535 | 1.32E-02 |
| regulation of transcription from RNA polymerase II promoter | 31 | 2.2286 12509 | 1.33E-02 |
| protein autophosphorylation | 19 | 1.3659 23796 | 1.38E-02 |
| transcription, DNA-templated | 59 | 4.2415 5284 | 1.47E-02 |
| positive regulation of transforming growth factor beta receptor signaling pathway | 5 | 0.3594 5363 | 1.56E-02 |
| protein phosphorylation | 17 | 1.2221 42344 | 1.57E-02 |
| DNA unwinding involved in DNA replication | 4 | 0.2875 62904 | 1.85E-02 |
| maintenance of DNA methylation | 4 | 0.2875 62904 | 1.85E-02 |
| regulation of attachment of spindle microtubules to kinetochore | 4 | 0.2875 62904 | 1.85E-02 |
| negative regulation of protein localization to nucleus | 4 | 0.2875 62904 | 1.85E-02 |
| mitotic spindle assembly | 7 | 0.5032 35083 | 1.87E-02 |
| RNA secondary structure unwinding | 9 | 0.6470 16535 | 1.93E-02 |
| positive regulation of protein ubiquitination | 9 | 0.6470 16535 | 1.93E-02 |
| intracellular signal transduction | 37 | 2.6599 56866 | 2.09E-02 |
| somatic stem cell population maintenance | 8 | 0.5751 25809 | 2.37E-02 |
| regulation of cell cycle | 12 | 0.8626 88713 | 2.41E-02 |
| positive regulation of cytokinesis | 6 | 0.4313 | 2.76E-02 |

| | | | |
|---|---|---|---|
| | | 44357 | |
| centriole replication | 4 | 0.2875 62904 | 2.99E-02 |
| protein localization to kinetochore | 4 | 0.2875 62904 | 2.99E-02 |
| chromosome organization | 5 | 0.3594 5363 | 3.09E-02 |
| replication fork processing | 5 | 0.3594 5363 | 3.09E-02 |
| negative regulation of proteasomal ubiquitin-dependent protein catabolic process | 5 | 0.3594 5363 | 3.09E-02 |
| B cell apoptotic process | 3 | 0.2156 72178 | 3.13E-02 |
| double-strand break repair via break-induced replication | 3 | 0.2156 72178 | 3.13E-02 |
| negative regulation of DNA damage checkpoint | 3 | 0.2156 72178 | 3.13E-02 |
| DNA replication, removal of RNA primer | 3 | 0.2156 72178 | 3.13E-02 |
| cleavage involved in rRNA processing | 3 | 0.2156 72178 | 3.13E-02 |
| PML body organization | 3 | 0.2156 72178 | 3.13E-02 |
| protein autoubiquitination | 9 | 0.6470 16535 | 3.17E-02 |
| cilium assembly | 15 | 1.0783 60891 | 3.32E-02 |
| thymus development | 8 | 0.5751 25809 | 3.39E-02 |
| protein localization to nucleus | 6 | 0.4313 44357 | 3.50E-02 |
| stem cell population maintenance | 9 | 0.6470 16535 | 3.68E-02 |
| positive regulation of NF-kappaB transcription factor activity | 12 | 0.8626 88713 | 3.89E-02 |
| regulation of Rho protein signal transduction | 11 | 0.7907 97987 | 3.89E-02 |
| activation of GTPase activity | 11 | 0.7907 97987 | 3.89E-02 |
| mitotic spindle organization | 5 | 0.3594 5363 | 4.10E-02 |
| DNA replication-dependent nucleosome assembly | 5 | 0.3594 5363 | 4.10E-02 |
| negative regulation of cell proliferation | 27 | 1.9410 49605 | 4.17E-02 |
| forebrain development | 7 | 0.5032 35083 | 4.24E-02 |
| DNA duplex unwinding | 6 | 0.4313 44357 | 4.35E-02 |
| positive regulation of transcription, DNA-templated | 30 | 2.1567 21783 | 4.41E-02 |
| protein localization to pre-autophagosomal structure | 4 | 0.2875 62904 | 4.42E-02 |
| mesoderm development | 4 | 0.2875 | 4.42E-02 |

| | | | |
|---|---|---|---|
| | | 62904 | |
| protein desumoylation | 4 | 0.2875 62904 | 4.42E-02 |
| regulation of mitotic nuclear division | 4 | 0.2875 62904 | 4.42E-02 |

**Significantly detected KEGG pathways based on up-regulated DEGs between oocyte and zygote (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| Cell cycle | 37 | 2.6599 56866 | 1.83E-11 |
| DNA replication | 18 | 1.2940 3307 | 1.99E-10 |
| Mismatch repair | 11 | 0.7907 97987 | 5.69E-06 |
| Ribosome biogenesis in eukaryotes | 20 | 1.4378 14522 | 1.06E-05 |
| Fanconi anemia pathway | 15 | 1.0783 60891 | 1.34E-04 |
| RNA transport | 27 | 1.9410 49605 | 3.23E-04 |
| Spliceosome | 23 | 1.6534 867 | 4.23E-04 |
| Oocyte meiosis | 19 | 1.3659 23796 | 1.86E-03 |
| Regulation of actin cytoskeleton | 30 | 2.1567 21783 | 3.65E-03 |
| Dorso-ventral axis formation | 8 | 0.5751 25809 | 3.89E-03 |
| Base excision repair | 9 | 0.6470 16535 | 4.33E-03 |
| p53 signaling pathway | 14 | 1.0064 70165 | 5.37E-03 |
| Nucleotide excision repair | 10 | 0.7189 07261 | 6.65E-03 |
| ErbB signaling pathway | 16 | 1.1502 51618 | 7.00E-03 |
| RNA degradation | 14 | 1.0064 70165 | 9.35E-03 |
| Notch signaling pathway | 10 | 0.7189 07261 | 2.48E-02 |
| Insulin signaling pathway | 19 | 1.3659 23796 | 3.04E-02 |
| Homologous recombination | 7 | 0.5032 35083 | 3.06E-02 |
| Wnt signaling pathway | 19 | 1.3659 23796 | 4.10E-02 |
| Hedgehog signaling pathway | 6 | 0.4313 44357 | 4.86E-02 |

**Table 6-7. Significantly detected SMART domain in three patterns of up-regulated TFs after fertilization**

**Up-regulated TFs that not significantly differ at EGK.X (enrichment test P-value < 0.05)**

| SMART term | # of genes | % | P-value |
|---|---|---|---|
| ZnF_C2H2 | 5 | 19.23076923 | 3.42E-03 |
| HMG | 3 | 11.53846154 | 6.22E-03 |

**Up-regulated TFs that re-upregulated until EGK.X (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| ZnF_C2H2 | 7 | 25.92592593 | 5.13E-05 |
| SM01372 | 3 | 11.11111111 | 3.55E-04 |
| SAM_PNT | 2 | 7.407407407 | 3.26E-02 |
| ZnF_GATA | 2 | 7.407407407 | 3.26E-02 |
| TBOX | 2 | 7.407407407 | 4.67E-02 |

**Up-regulated TFs that down-regulated until EGK.X (enrichment test P-value < 0.05)**

| KEGG term | # of genes | % | P-value |
|---|---|---|---|
| ZnF_C2H2 | 15 | 20.54794521 | 1.17E-09 |
| HLH | 9 | 12.32876712 | 8.27E-07 |
| HMG | 4 | 5.479452055 | 4.21E-03 |
| BRLZ | 4 | 5.479452055 | 4.21E-03 |
| SANT | 3 | 4.109589041 | 2.16E-02 |
| BTB | 5 | 6.849315068 | 3.15E-02 |

**Table 6-8. Detected SNPs in each chromosome on the WGS data**

| Chr. | Length | Variants | Chr. | Length | Variants |
|---|---|---|---|---|---|
| 1 | 195,276,750 | 2,109,432 | 17 | 10,454,150 | 111,208 |
| 2 | 148,809,762 | 1,601,156 | 18 | 11,219,875 | 117,511 |
| 3 | 110,447,801 | 1,199,589 | 19 | 9,983,394 | 104,435 |
| 4 | 90,216,835 | 1,003,785 | 20 | 14,302,601 | 148,040 |
| 5 | 59,580,361 | 633,183 | 21 | 6,802,778 | 76,256 |
| 6 | 34,951,654 | 450,107 | 22 | 4,081,097 | 23,779 |
| 7 | 36,245,040 | 414,100 | 23 | 5,723,239 | 63,649 |
| 8 | 28,767,244 | 310,967 | 24 | 6,323,281 | 73,649 |
| 9 | 23,441,680 | 285,875 | 25 | 2,191,139 | 16,869 |
| 10 | 19,911,089 | 224,814 | 26 | 5,329,985 | 60,495 |
| 11 | 19,401,079 | 199,959 | 27 | 5,209,285 | 51,527 |
| 12 | 19,897,011 | 236,818 | 28 | 4,742,627 | 47,640 |
| 13 | 17,760,035 | 200,916 | W | 1,248,174 | 65 |
| 14 | 15,161,805 | 167,141 | Z | 82,363,669 | 471,462 |
| 15 | 12,656,803 | 121,225 | Total | 1,003,035,513 | 10,529,469 |
| 16 | 535,270 | 3,817 | | | |

**Table 6-9. Detected SNPs in each chromosome on the WTS data**

| Chr. | Length | Variants | Chr. | Length | Variants |
|---|---|---|---|---|---|
| 1 | 195,276,750 | 39,875 | 17 | 10,454,150 | 5,283 |
| 2 | 148,809,762 | 27,170 | 18 | 11,219,875 | 5,386 |
| 3 | 110,447,801 | 25,102 | 19 | 9,983,394 | 5,179 |
| 4 | 90,216,835 | 23,099 | 20 | 14,302,601 | 5,509 |
| 5 | 59,580,361 | 20,014 | 21 | 6,802,778 | 4,129 |
| 6 | 34,951,654 | 10,772 | 22 | 4,081,097 | 706 |
| 7 | 36,245,040 | 11,061 | 23 | 5,723,239 | 2,781 |
| 8 | 28,767,244 | 10,114 | 24 | 6,323,281 | 2,620 |
| 9 | 23,441,680 | 8,168 | 25 | 2,191,139 | 1,086 |
| 10 | 19,911,089 | 7,792 | 26 | 5,329,985 | 3,488 |
| 11 | 19,401,079 | 6,046 | 27 | 5,209,285 | 2,441 |
| 12 | 19,897,011 | 6,346 | 28 | 4,742,627 | 3,630 |
| 13 | 17,760,035 | 5,786 | W | 1,248,174 | 92 |
| 14 | 15,161,805 | 8,312 | Z | 82,363,669 | 7,019 |
| 15 | 12,656,803 | 6,293 | Total | 1,003,035,513 | 265,788 |
| 16 | 535,270 | 489 | | | |

**Table 6-10. Primers used for the exon-intron RT-PCR**

| Gene name | Wave | Location | Primer sequences | |
|---|---|---|---|---|
| | | | Forward (5'→3') | Reverse (5'→3') |
| DLX6 | 1st | Exon 3 - Intron 3 | CCACACCAAGACACAATGCA | CTCCTTTCCTCTCCCCGTTT |
| GATA2 | 1st | Exon 3 - Intron 3 | AGGGATGAAGATGGAGAGCG | TTTGTACGTGCGATCTGCTG |
| ZIC4 | 1st | Intron 2 - Exon 3 | CTCTCCGGTCACACAGAGG | CCAGGTTCACGTTCAGGTTC |
| WNT11 | 2nd | Exon 4 - Intron 4 | AGACATTGCACTGGACCTCA | TTTAAGGCTGCCTGGAAGGA |
| WNT3A | 2nd | Exon 2 - Intron 2 | GGGGTGAAGATTGGGATCCA | ACCTTACACCAGAGAGCCAC |
| C8ORF22 | 2nd | Intron 1 - Exon 2 | GTGCACCCAACCTGGTTAAG | TGGATTCCGAGTTCTGCCTT |

**Table 6-11. Primers used for the validation of allelic expression**

| Gene name | SNP ID | Primer sequences | |
|---|---|---|---|
| | | Forward (5'→3') | Reverse (5'→3') |
| MAP7D1 | rs735616631, rs737091869, rs317162724, rs312431159 | TCACGCTGGAAATGCTTCTC | AGAAGAAGGAGAAGGAGCGG |
| | rs732500851 | GCGGTTTCTACTGGGAGAGT | CTGCATCACCATCTTCAGCC |
| ESCO1 | rs314014865, rs312569361, rs317102766, rs315302113 | TGCTGCTAAGAACGGGAGAT | ACACGCATGACTACAAGGTC |
| CCNB3 | rs731100154, rs739402820 | CCTCATCGCCTCCAAATTCG | ATCATGTCCATGGGGGTTGT |
| SYTL4 | rs313237661, rs317531504, rs315733912, rs317004903, rs317743001 | GGACAGACTCTCCTGCAGAA | TCGTTGTACAGCGGGTTTAC |
| GRHL1 | rs15435942, rs313477995, rs316662644 | ACAAGGTTCCAAGGGAGAGG | GGATTCCTTCACTGCTGCTG |
| LLGL1 | rs314448475 | CAAAGCGTGCCCGATGATTA | GGTTCTCAAAGCCTCCCTCT |
| | rs315560932, rs735260738 | CTGGTGATGGAGCTGAGTGA | AGGAGATGAGCATGGAGTGG |

## 4. Discussion

Here, we provided the profound analysis of transcripts by 1st wave in chicken using a multi-omics approach. Our results indicated that 1st transcription after fertilization exclusively activated maternal alleles in chicken (Figure 6-9C). After MZT, however, most expressed genes were from the paternal and maternal genomes. Functionally, transcripts by 1st wave were involved in asymmetric rapid cellularization and fundamental regulation of further development (Figure 6-9C).

We speculate that this evolved by necessity in animals with physiological polyspermy (Figure 6-11). Polyspermy is usually embryonic lethal because of pathological mitosis (Snook et al., 2011). However, the polyspermic mechanism of fertilization in several species, including avian and amphibian species, is essential for fertilization and early development due to the need for a number of sperm to activate large eggs (Iwao, 2012). In addition to pathological mitosis in monospermic animals, the transcriptional activation in polyspermic embryos is greatly stimulated by the supernumerary sperm pronucleus compared with monospermic embryos in sea urchin (Poccia et al., 1985). The disproportionate genome contribution could result in an excessive amount of global genetic material. The total polyspermy number reportedly varied from 29 to 164,000 per egg among birds, and was positively correlated with egg size (Birkhead et al., 1994). Even, a high degree of variation in sperm number was observed in chicken, ranging from 1–10 to more than 1,000 per EGK.I-III embryo naturally, and depended on the environment such as sperm-limited condition (Lee et al., 2013, Hemmings and Birkhead, 2015). Therefore, polyspermic animals may have evolved means of inhibiting the activation of paternal PN to control gene expression levels from the 1st transcriptional activation.

Moreover, in monospermic mammals, *in vivo* polyspermy incidence is about 1% to 2% (Gardner and Evans, 2006). But in the case of *in vitro* fertilization, the frequency of polyspermy increases with sperm concentration (McAvey et al., 2002), indicating polyspermy mechanism could be important to reproductive technology. Thus, the pathological polyspermy in mammalian fertilization also should be investigated from the perspective of controlling paternal transcription, refer to allelic expression in avian zygotes.

In addition, genetic stability of the fertilized embryo may be an extra consideration in avian species, as well as quantitative gene expression due to polyspermy. Individual sperm provide genomic diversity (Wang et al., 2012) but result in genomic instability if different types of transcript are expressed by various sperm nuclei. To inhibit inattentive activation of the paternal genome, there may be avian-specific mechanisms of allelic histone modification after fertilization, regardless of genomic imprinting which is absent in chicken (Fresard et al., 2014).

Taken together, our results defined the definite two waves of transcriptional activation after fertilization and EGK.V. Intriguingly, maternal genome was exclusively activated after fertilization, regardless of the presence of haploid genomes in male PN and supernumerary sperm in the egg. In addition, this maternal-derived transcripts regulated the cleavage stages and was changed to biallelic expression in later stages. This study indicate that maternally derived 1$^{st}$ expression alone is essential for early development and evolutionary outcome in avian species.

# CHAPTER 7

# The Avian-Specific Small Heat Shock Protein HSP25 is a Constitutive Protector against Environmental Stresses during Blastoderm Dormancy

## 1. Introduction

Heat shock protein (HSP) levels increase in response to various cellular stresses and they function as molecular chaperones which bind to and inhibit irreversible protein aggregation or misfolding under stressful conditions (Richter et al., 2010). Among HSP families, members of the small heat shock protein (sHSP) family range in size from 12 to 42 kDa, and possess highly variable N-terminal and C-terminal regions and conserved α-crystallin domains (Basha et al., 2012). Monomers of sHSP can interact and bind themselves via the α-crystallin domain to form dimers or higher oligomers, assisted by the N- and C-terminal regions. Unlike other HSPs, sHSPs function as holdases in the absence of ATP and can bind various protein substrates, thereby contributing to cell survival(Mymrikov et al., 2011). This ATP-independent holdase function of sHSPs is especially important when the ATP concentration is low or limited.

Induction of sHSP expression is stimulated by stress, but may also be under developmental control or regulated in a cell- and tissue-specific manner (Morrow and Tanguay, 2012, Saretzki et al., 2008). During early gastrulation in *Danio rerio*, *heat shock protein family B (small) member 1* (*hspb1*) gene is expressed transiently in the developing myotome, lens, and presumptive brain(Mao and Shelden, 2006). Constitutive expression of the *hsp30* genes has been detected in the cement gland of early and mid-tailbud *Xenopus laevis* embryos(Lang et al., 1999). In mouse pre-implantation embryos, *Hspb1* mRNA is induced by zygotic genome activation at the two-cell stage, is subsequently decreased at the four-cell stage, and is re-upregulated at the morula stage, with the highest expression in the blastocyst (Kim et al., 2002). Additionally, mouse embryonic stem cells show a unique stress-resistant gene

expression signature, including *Hspb1*, which becomes downregulated during embryoid body formation (Saretzki et al., 2004).

In particular, sHSP genes are expressed specifically in the dauer stage of *Caenorhabditis elegans* and in the cyst of *Artemia franciscana*, which are so-called diapause states described in many reports as stress-tolerant and developmental arrest (Ludewig et al., 2004, MacRae, 2016). The *C. elegans* dauer stage with DAF-16 activity undergoes marked induction of several sHSPs, including *hsp16.1*, *hsp16.49*, *hsp-12.6*, *hsp-12.3*, *hsp-20*, and *sip-1* (Hsu et al., 2003, McElwee et al., 2004). In the case of *A. franciscana*, *p26*, *ArHsp21*, and *ArHsp22* are expressed specifically in diapause, and reach a peak of expression in the cyst, while they are not detected during development (Villeneuve et al., 2006, Qiu and MacRae, 2008b, Qiu and Macrae, 2008a, King et al., 2013). Moreover, a lack of p26 in the cyst resulted in the spontaneous termination of diapause (King and MacRae, 2012).

Eleven small heat shock protein genes from chicken are entered in the GenBank database, including the HSP30 subfamily, which is restricted to oviparous animals (Katoh et al., 2004). Chicken sHSPs induced by heat shock and chemical stresses have been studied and observed in somatic fibroblast cells (Katoh et al., 2004) and adult tissues (Kawazoe et al., 1999, Wang et al., 2013, Luo et al., 2014, Wang et al., 2015). In avian species, although embryonic diapause has not been reported, developmental arrest, so-called cold torpor, of the embryos of the earlier eggs in a single clutch is common at the beginning of incubation, until all eggs in the clutch have been laid (Welty, 1982). Bloom et al. reported that chicken Eyal-Giladi and Kochav (EGK) stage X blastoderms after oviposition (Eyal-Giladi and Kochav, 1976) could survive 3-4 weeks in a dormant state under conditions of cold storage and that

this involved expressing *HSP70* and anti-apoptotic B-cell CLL/lymphoma *(BCL*) *family* genes (Bloom et al., 1998). However, the expression and functions of small HSPs in the stress-tolerant avian blastoderm have yet to be investigated.

In this study, we first identified the expression and function of sHSPs in the blastoderm to assess which sHSP was associated with the stress-resistant characteristics of blastoderm cells in chickens. We also demonstrated that avian *HSP25* was important in the protection of future embryonic cells in the blastoderm.

## 2. Materials and methods

### *Experimental animals and animal care*

The care and experimental use of chickens was approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1). Chickens were maintained according to a standard management program at the University Animal Farm, Seoul National University, Korea. The procedures for animal management, reproduction, and embryo manipulation adhered to the standard operating protocols of our laboratory.

### *Multiple sequence alignment, pair-wise comparison, and phylogenetic analysis*

Amino acid sequences of small heat shock proteins containing the α-crystallin domain HSPB9-like from the NCBI database that were analyzed included chicken HSP25, HSP30C-like (NP_001010842, XP_003642880), Japanese quail HSP25, HSP30C-like (XP_015741136, XP_015741135), collared flycatcher HSP30C-like (XP_005059698), Tibetan ground-tit HSP30C-like (XP_005531515), zebra finch HSP25-like, HSP30C-like (NP_001232665, XP_004174712), rock pigeon HSPB11-like, HSP30C (XP_005513649, XP_005513650), African clawed frog HSP30C, HSP30D (NP_001165977, NP_001165976), Atlantic salmon HSP30 (NP_001134440), human HSPB9 (NP_149971), and mouse HSPB9 (NP_083583). For pair-wise comparisons, multiple sequence alignments, and phylogenetic trees, the amino acid sequences of the small heat shock proteins above were aligned using the Geneious software (ver. 6.0.4; Auckland, New Zealand) with default penalties for gaps and the protein weight matrix of BLOSUM (blocks substitution matrix). A phylogenetic tree was reconstructed using a neighbour-joining

method.

***Sample preparation***

Blastoderms were obtained within 6 h after oviposition from White Leghorn chickens (WL), Japanese quail (JQ), and zebra finch (ZF). WL, JQ, and ZF eggs were incubated with intermittent rocking at 37.5°C under 60-70% relative humidity. Chicken blastoderm cells were collected by gentle dissociation of EGK stage X blastoderms from WL eggs(Eyal-Giladi and Kochav, 1976). WL eggs were stored at 16°C under 70-80% relative humidity for 2 weeks. The egg-laying times of the WL hens were recorded and intrauterine eggs from EGK stages III-VIII were harvested using an abdominal massage technique(Lee et al., 2013). To collect oocytes and zygotes, WL hens were sacrificed and the follicles were collected. Chicken embryonic fibroblasts (CEFs), quail embryonic fibroblasts (QEFs), and zebra finch embryonic fibroblasts (ZEFs) were collected by dissociating the embryonic body of Hamburger and Hamilton (HH) stage 28 in 0.05% trypsin-EDTA (GIBCO Invitrogen, Grand Island, NY, USA) at 37°C for 10 min(Hamburger and Hamilton, 1951). Cells were then cultured in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific, Inc., Waltham, MA, USA) containing 10% FBS and 1% antibiotic-antimycotic (Invitrogen) in a 5% $CO_2$ atmosphere at 37°C.

***RT-PCR and quantitative real-time PCR analysis***

Total RNA was isolated using the Trizol Reagent (Invitrogen, Carlsbad, CA), according to the manufacturer's protocol. For RT-PCR and quantitative real-time PCR analysis of mRNAs, total RNA (1 μg) was used as template for cDNA synthesis using the SuperScript III First-Strand Synthesis

System (Invitrogen). The cDNA was serially diluted five-fold and was equalized quantitatively for PCR amplification. Primers for real-time PCR of each gene transcript were designed using the program Primer3 (Table 7-1; http://frodo.wi.mit.edu/). RT-PCR was performed with an initial incubation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 59°C for 30 s, and 72°C for 30 s. The reaction was terminated after a final incubation at 72°C for 5 min. Real-time PCR analysis was performed using a CFX96 real-time PCR detection system with a C1000 thermal cycler (Bio-Rad Laboratories, Hercules, CA, USA). The qRT-PCR conditions were 95°C for 3 min followed by 40 cycles of 95°C for 30 s, 59°C for 30 s, and 72°C for 30 s. Melting curve profiles were analyzed for amplicons. Each test sample was run in triplicate. The relative quantification of gene expression was analyzed with the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

***Chicken blastoderm cell culture***

Stage X blastoderm cells were prepared as described previously (Lee et al., 2011). Briefly, blastoderm cells were prepared from the pellucida area of WL embryos at stage X and dissociated mechanically into single cells. Single stage X cells were centrifuged (850 rpm, 3 min) and washed twice in PBS. Blastoderm cells were cultured according to previous descriptions with minor modifications (Nakanoh et al., 2013). Cells were cultured in N2B27/2i medium containing DMEM/F-12 (Gibco), Neurobasal (Gibco), 55 mM β-mercaptoethanol (Gibco), 200 mM L-glutamine (Gibco), N2-Supplement (100×, Gibco), and B27 supplement-vitamin A (50×, Gibco). The medium was supplemented with two inhibitors, 3 μM CHIR99021 and 1 μM PD0325901 (Stemgent, San Diego, CA, USA).

***Transfection and chemical treatment of chicken blastoderm cells***

*HSP25*-specific siRNAs were designed and purchased from Bioneer Corporation (Table 7-2; Daejeon, Korea). For the transfection of siRNAs into the cultured blastoderm cells, Lipofectamine RNAiMAX (Invitrogen) was used according to the manufacturer's protocol for 48 h. For the *in vitro* induction of apoptosis, reactive oxygen species (ROS) production, and autophagy, cells were treated with mitomycin C (Sigma-Aldrich, St. Louis, MO, USA), $H_2O_2$ (Sigma), and MG132 (Sigma). Specifically, 0, 30, and 60 μM mitomycin C was added for 4 h; 0, 100, and 200 μM $H_2O_2$ was added for 4 h; and 0, 5, and 10 μM MG132 was added for 24 h. After each treatment, cells were harvested and analyzed.

***TUNEL assay***

Cells were washed and concentrated on glass slides. After fixation in 2% paraformaldehyde for 15 min, the cells were incubated in a permeablization solution (0.1% Triton X-100 in PBS) for 10 min. Apoptotic cells were identified using an In Situ Cell Death Detection Kit and TMR red (Roche Applied Science, Basel, Switzerland) that stains apoptotic cells red. Cells were counterstained with DAPI, mounted, and analyzed under a fluorescence microscope (TU-80, Nikon).

***Flow cytometry***

To examine apoptotic cell death, annexin V/propidium iodide (PI) (Thermo Fisher Scientific) staining was performed according to manufacturer's protocol and cells were analyzed by flow cytometry (FACSCalibur, BD Biosciences, San Jose, CA, USA). Annexin V and PI double-negative cells were considered viable, annexin V-positive and PI-negative (early apoptotic) cells and annexin V- and PI-positive (late apoptotic)

cells were considered apoptotic, and PI only-positive cells were considered necrotic. To assess cellular ROS, a 2',7'-dichlorofluorescein diacetate (DCFDA) assay (Abcam, Cambridge, UK) was performed according to the manufacturer's protocol and analyzed with the FACSCalibur. To assay autophagy activity, cells were labelled with anti-light chain 3 (LC3) (Abcam) followed by staining with Alexa488-conjugated goat anti-mouse IgG antibodies. Next, cells were analyzed with FACSCalibur and data were analyzed with the FlowJo software (ver. 7.6.5; Tree Star, Ashland, OR, USA).

***Statistical analyses***

Significant differences between groups were examined statistically using Student's *t*-test and one-way ANOVA. A $P$ value < 0.05 was considered to indicate statistical significance (***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$).

## 3. Results

### *Expression profiling of sHSPs in chicken stage X blastoderm*

To identify the expression of sHSPs in chicken stage X blastoderm, 11 *sHSP* genes from the GenBank database were examined by RT-PCR and qRT-PCR using cDNA from CEFs and stage X blastoderms. *HSP25* and *HSP30CL* were detected specifically in stage X blastoderm. However, *HSPB1* was expressed specifically in CEFs and *HSPB8* was expressed more abundantly in CEFs (Figure 7-1A). Other small heat shock genes were not detected in either. Furthermore, qRT-PCR analysis showed that *HSP25* and *HSP30CL* expression levels were 72.2-fold and 15.7-fold higher in stage X blastoderms than CEFs, respectively (Figure 7-1B), while *HSPB1* and *HSPB8* expression levels were 6-fold and 1.4-fold lower in stage X blastoderms than CEFs, respectively (Figure 7-1C).

### *Expression of chicken stage X blastoderm-specific HSP30 subfamily members during intrauterine development, egg incubation, and egg storage*

As shown above, chicken *HSP25* and *HSP30CL* were expressed specifically in stage X blastoderms. Next, we investigated the expression patterns during intrauterine and early development and egg storage. First, we examined the expression patterns of the HSP30 subfamily during chicken intrauterine development. Expression of both genes was detected at low levels until EGK.III at which point zygotic gene activation (ZGA) occurred. In the case of *HSP25*, after EGK.VIII, expression increased sharply only shortly before oviposition, whereas, *HSP30CL* expression increased after EGK.III and EGK.VIII (Figure 7-2A). Furthermore, after a 4 h incubation of eggs at 37.5°C to promote embryonic development to the pre-streak stage(Eyal-Giladi

and Kochav, 1976), expression of *HSP25* decreased significantly. However, *HSP30CL* expression in the 4-h-incubated embryo was significantly higher than in stage X blastoderm (Figure 7-2A). To examine *HSP30* subfamily expression dynamics due to increased stress in the stage X blastoderm, we also quantified transcripts during egg storage at 16°C. Expression of both *HSP25* and *HSP30CL* was upregulated gradually during egg storage. However, *HSP25* mRNA expression increased significantly after 5 days of storage versus the unstored blastoderm and this was maintained until 14 days, whereas upregulation of *HSP30CL* mRNA expression was not significant (Figure 7-2B).

***Multiple sequence alignment, phylogenetic analysis, and expression in other avian blastoderms of sHSPs containing the α-crystallin domain HSPB9-like***

mRNA and protein sequences of chicken HSP25 were obtained from the NCBI *Gallus gallus* genome database. The *HSP25* mRNA sequence contains an open reading frame of 582 base pairs with no intron sequence and encodes 194 amino acids, including the α-crystallin domain HSPB9-like. Next, we aligned the protein sequences containing the α-crystallin domain HSPB9-like from chicken HSP25, HSP30C-like, Japanese quail HSP25, HSP30C-like, collared flycatcher HSP30C-like, Tibetan ground-tit HSP30C-like, zebra finch HSP25-like, HSP30C-like, rock pigeon HSPB11-like, HSP30C, African clawed frog HSP30C, HSP30D Atlantic salmon HSP30, human HSPB9, and mouse HSPB9 (Figure 7-3A). Pair-wise comparisons of sHSPs showed high degrees of homology among avian species including the α-crystallin domain, but they differed from other vertebrate sequences. Also, in the phylogenetic analysis, and the degree of similarity to chicken HSP25,

they separated into two subfamilies, forming a branch separated from other vertebrates, one including chicken HSP25, Japanese quail HSP25, collared flycatcher HSP30C-like, Tibetan ground-tit HSP30C-like, zebra finch HSP25-like, and rock pigeon HSPB11-like, and another including chicken HSP30C-like, Japanese quail HSP30C-like, zebra finch HSP30C-like, and pigeon HSP30C due to N-terminal variations between the subfamilies (Figure 7-3B, C). To identify the expression patterns of *HSP25* and *HSP30CL* in other avian species, we compared QEFs, ZEFs, and blastoderms at oviposition by RT-PCR. Similar to chicken, expression of the Japanese quail and zebra finch *HSP25* and *HSP30CL* genes was detected specifically in blastoderms (Figure 7-3D). Furthermore, qRT-PCR analysis showed that *HSP25* mRNA expression increased significantly after 7 days of storage versus the unstored blastoderm in quail and zebra finch (Figure 7-3E), similar to the case of chicken embryo.

***Effect of HSP25 knockdown in chicken blastoderm cells***

To investigate chicken *HSP25* function, we designed two candidate siRNAs, siRNA-296 and siRNA-497, and transfected them into blastoderm cells *in vitro*. At 48 h after transfection, the *HSP25* transcript level was analyzed using real-time PCR. As shown in Figure 7-4A, siRNA-296 was the most efficient knockdown probe for *HSP25*, showing 68% suppression versus the control ($P < 0.01$). This siRNA was selected for further knockdown experiments of chicken *HSP25*. First of all, we examined the role of chicken *HSP25* in stem cell property, like as mouse ESCs (Saretzki et al., 2004, Saretzki et al., 2008). After knockdown of *HSP25* in chicken blastoderm cells, we examined the expression of pluripotency markers, including *nanog homeobox* (*NANOG*), *POU domain class 5 transcription factor 3* (*POUV*),

*SRY (sex determining region Y)-box 2* (*SOX2*), and *cripto, FRL-1, cryptic family 1B* (*CRIPTO*). Compared with the control, knockdown of *HSP25* in blastoderm cells decreased the expression of all pluripotency genes significantly ($P < 0.05$; Figure 7-4B). Furthermore, we performed a TUNEL assay to examine the effects of *HSP25* knockdown on apoptosis in chicken blastoderm cells. Apoptotic signals were not detected in the control samples, but were strongly induced in most blastoderm cells after *HSP25* knockdown (Figure 7-4C).

***Anti-apoptotic function of HSP25 in chicken blastoderm cells***

Next, we investigated chicken *HSP25* function in blastoderm cell survival under various stress conditions. First, to study the anti-apoptotic function of chicken *HSP25*, we assessed apoptotic cell death in blastoderm cells after *HSP25* knockdown followed by mitomycin C treatment with annexin V/PI and flow cytometry analysis (Figure 7-5A). After mitomycin C treatment on blastoderm cells, the proportion of live cells was decreased (annexin $V^-$ and $PI^-$, Figure 7-5B) and apoptotic cells (annexin $V^+$ only + annexin $V^+$ and $PI^+$, Figure 7-5C) was increased significantly. Then, compared with the control, *HSP25* knockdown decreased the proportion of live cells significantly when treated with mitomycin C at 30 μM and 60 μM ($P < 0.001$ and $< 0.01$, respectively; Figure 7-5B). Also, Figure 7-5C shows that the repression of *HSP25* increased the proportion of apoptotic cells significantly (30 μM, $P < 0.05$ and 60 μM, $P < 0.01$). Correspondingly, expression of anti-apoptotic genes, such as *BCL2* and *BCL2L1*, was decreased significantly by *HSP25* knockdown followed by 60 μM mitomycin C treatment in chicken blastoderm cells (Figure 7-5D)

***Anti-oxidant and anti-necrotic effects of HSP25 in chicken blastoderm cells***

Chicken blastoderm cells were treated with $H_2O_2$ to create oxidative stress after *HSP25* knockdown, then a DCFDA assay was performed with flow cytometry to measure ROS in blastoderm cells (Figure 7-6A). The DCFDA analysis showed that ROS production was significantly higher after *HSP25* knockdown than the control with or without $H_2O_2$ treatment ($P < 0.05$; Figure 7-6B). Furthermore, we conducted a gene expression analysis of anti-oxidant genes by real-time PCR. Figure 7-6C shows that two anti-oxidant genes, *glutathione peroxidase 3* (*GPX3*) and *glutathione peroxidase 4* (*GPX4*), were downregulated after *HSP25* knockdown followed by 200 μM $H_2O_2$ treatment.

Generally, oxidative stress due to $H_2O_2$ promotes cell death in various cell types. Thus, we also used annexin V/PI and flow cytometry to investigate effects of *HSP25* on cell death with oxidative stress (Figure 7-7A). After knockdown of *HSP25*, double negative annexin V/PI cells (live cells; Figure 7-7B) were decreased significantly with $H_2O_2$ treatment (100 μM, $P < 0.01$ and 200 μM, $P < 0.05$). Notably, although the proportion of apoptotic cells (annexin $V^+$ only + annexin $V^+$ and $PI^+$) was slightly, but not significantly increased (Figure 7-7B), that of necrotic cells ($PI^+$ only) was increased significantly after *HSP25* knockdown followed by $H_2O_2$ treatment (100 μM, $P < 0.001$ and 200 μM, $P < 0.01$; Figure 7-7C).

***Pro-autophagic function of HSP25 in chicken blastoderm cells***

We examined the pro-autophagic function of chicken *HSP25* in blastoderm cells by siRNA knockdown and autophagy stimulation experiments. After 48 h of knockdown of chicken *HSP25*, we treated cells with MG132 at 5 and 10 μM to induce autophagy, then analysed the cells using flow cytometry for LC3 antibody-Alexa488 fluorescence (Figure 7-8A).

After MG132 treatment on blastoderm cells, the stimulation of autophagy was increased significantly (Figure 7-8B). Then, quantification of LC3-positive cells showed that *HSP25* knockdown decreased autophagy activation significantly under MG132 induction versus the control. However, there was no difference between knockdown and the control without MG132 induction (Figure 7-8B). Real-time PCR analysis of pro-autophagic genes was also performed. Figure 7-8C shows that expression levels of several pro-autophagic genes, such as *phosphatase and tensin homolog* (*PTEN*), *beclin 1, autophagy related* (*BECN1*), *UV radiation resistance associated* (*UVRAG*), *autophagy related 12* (*ATG12*), and *autophagy related 5* (*ATG5*), were downregulated after *HSP25* knockdown followed by 10 μM MG132-induced autophagy.

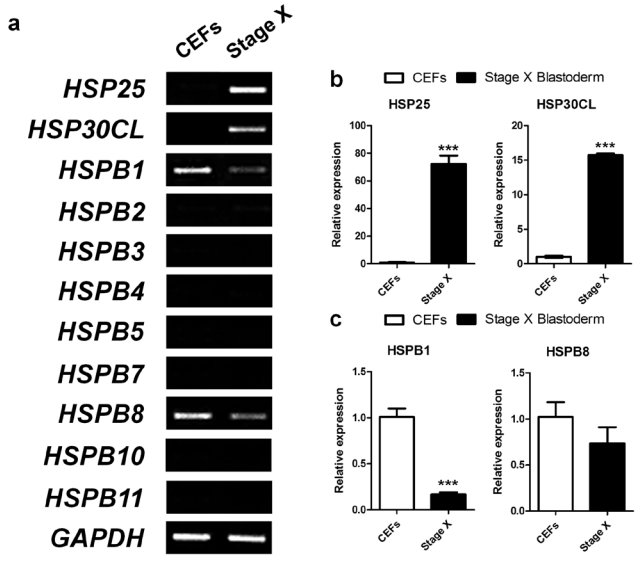


**Figure 7-1**. **Expression profiling of sHSPs.** Expression analysis of *sHSPs* in chicken embryonic fibroblasts (CEFs) and chicken stage X blastoderm (Stage X) by RT-PCR (A) and quantitative real-time PCR (B, C). *HSP25* and *HSP30CL* were detected specifically in stage X blastoderm (B). Real-time PCR was conducted in triplicate and normalised to expression of *GAPDH*. Significant differences between groups are indicated as ***$P < 0.001$. Error bars indicate the SE of triplicate analyses.

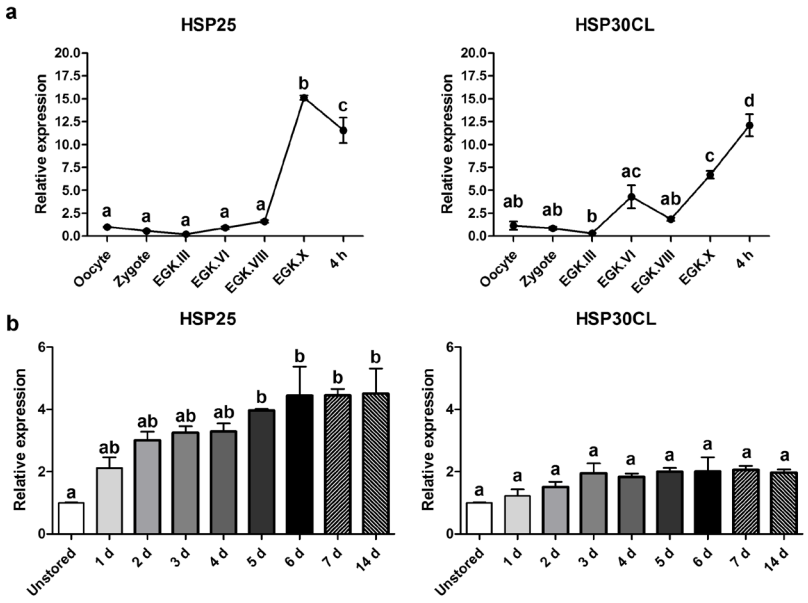


**Figure 7-2**. **Expression dynamics of chicken stage X blastoderm-specific HSP30 subfamily.** Quantitative expression analysis of blastoderm-specific chicken *HSP30 subfamily* during intrauterine and early development (A) and egg storage (B). Real-time PCR was conducted in triplicate and normalised to expression of *GAPDH* and *ACTB*. Significant differences between groups are indicated by different letters. Error bars indicate the SE of triplicate analyses. EGK, Eyal-Giladi and Kochav stage; 4 h, 4-h-incubated embryo at 37.5°C; d, days of storage at 16°C.

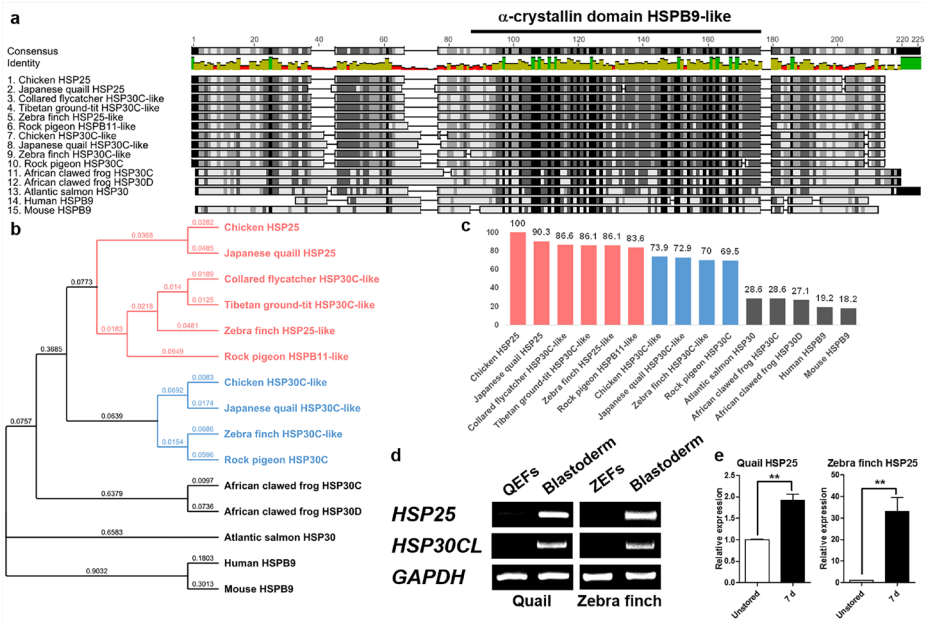


**Figure 7-3**. **Multiple alignment and phylogenetic trees of sHSPs containing the α-crystallin domain HSPB9-like.** Multiple sequence alignment (A), phylogenetic trees (B), and similarity to chicken HSP25 (C) of the amino acid sequences of sHSPs containing the α-crystallin domain HSPB9-like. Bold lined amino acids indicate the α-crystallin domain HSPB9-like. Avian HSP25 in red and HSP30CL in blue. Expression analysis of *HSP25* and *HSP30CL* in quail and zebra finch by RT-PCR (D) and during egg storage by qRT-PCR (E). Real-time PCR was conducted in triplicate and normalised to expression of *GAPDH*. Significant differences between groups are indicated as $^{**}P < 0.01$. Error bars indicate the SE of triplicate analyses. QEFs, quail embryonic fibroblasts; ZEF, zebra finch embryonic fibroblasts; Blastoderm, blastoderm at oviposition in each species; d, days of storage at 16°C.

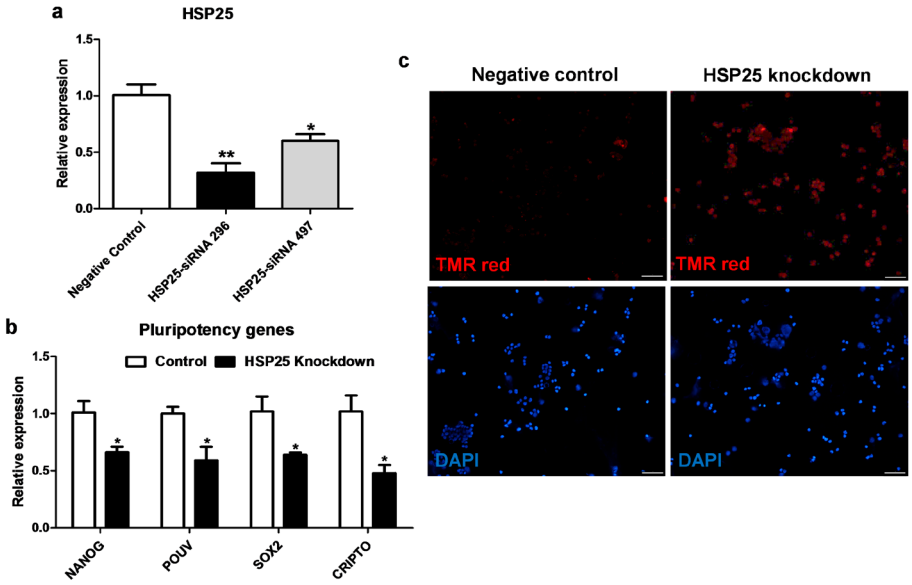


**Figure 7-4**. **Knockdown analysis of HSP25 in chicken blastoderm cells.** (A) Knockdown efficiency of *HSP25*-specific siRNAs in chicken blastoderm cells. Non-complementary sequences in the chicken genome were used as a control. (B) Relative expression analysis of pluripotency marker genes after *HSP25* knockdown. Real-time PCR was conducted in triplicate and normalised to the expression of *GAPDH*. Significant differences between groups are indicated as $^{**}P < 0.01$ and $^{*}P < 0.05$. Error bars indicate the SE of triplicate analyses. (C) TUNEL assay performed on chicken blastoderm cells after *HSP25* knockdown. Scale bars are 50 µm.

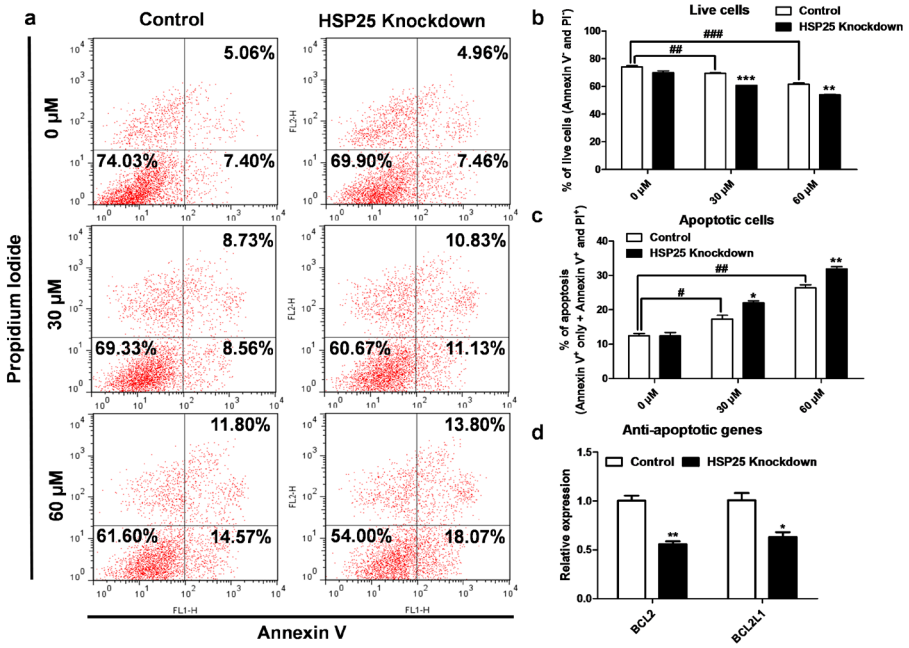


**Figure 7-5**. **Anti-apoptotic function of HSP25 in mitomycin C-treated chicken blastoderm cells.** (A) Annexin V/PI analysis by flow cytometry after *HSP25* knockdown followed by mitomycin C treatment (0, 30, or 60 μM). Quantitative analysis of double-negative (B) and double-positive (C) cells, indicating live and apoptotic cells, respectively. (D) Relative expression analysis of anti-apoptotic genes after *HSP25* knockdown followed by 60 μM mitomycin C treatment. Non-complementary sequences in the chicken genome were used as a control. Real-time PCR was conducted in triplicate and normalised to expression of *GAPDH*. ###$P < 0.001$ and ##$P < 0.01$ significance of mitomycin C treatment (30, or 60 μM) compared to 0 μM. Significant differences between control and *HSP25* knockdown are indicated as ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$. Error bars indicate the SE of triplicate analyses.

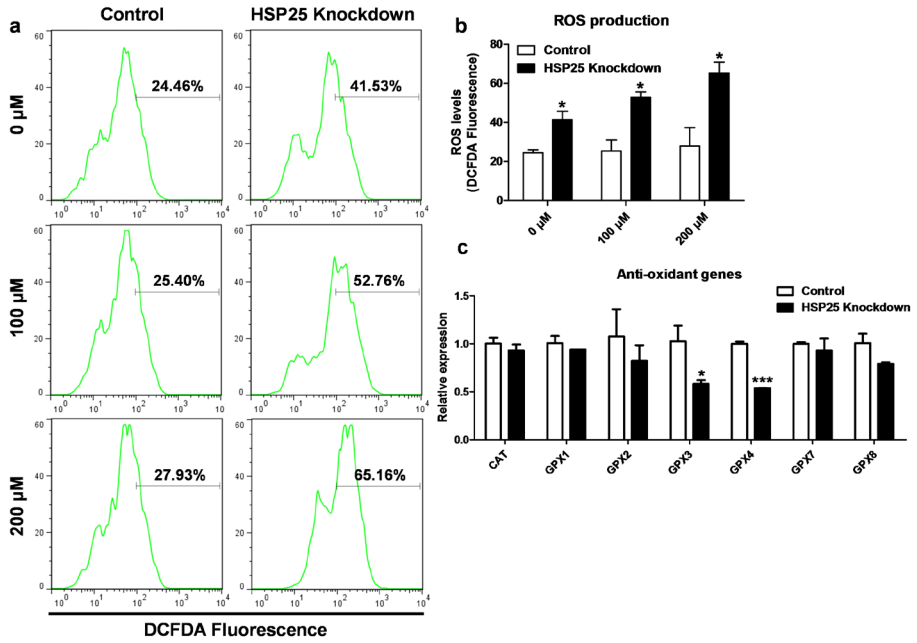


**Figure 7-6**. **Anti-oxidant effect of HSP25 in $H_2O_2$-treated chicken blastoderm cells.** (A) DCFDA analysis by flow cytometry after *HSP25* knockdown followed by $H_2O_2$ treatment (0, 100, or 200 μM). (B) Quantitative analysis of DCFDA-positive cells. (C) Relative expression analysis of anti-oxidant genes after HSP25 knockdown followed by 200 μM $H_2O_2$ treatment. Non-complementary sequences in the chicken genome were used as a control. Real-time PCR was conducted in triplicate and normalised to expression of *GAPDH*. Significant differences between control and *HSP25* knockdown are indicated as ***$P < 0.001$ and *$P < 0.05$. Error bars indicate the SE of triplicate analyses.

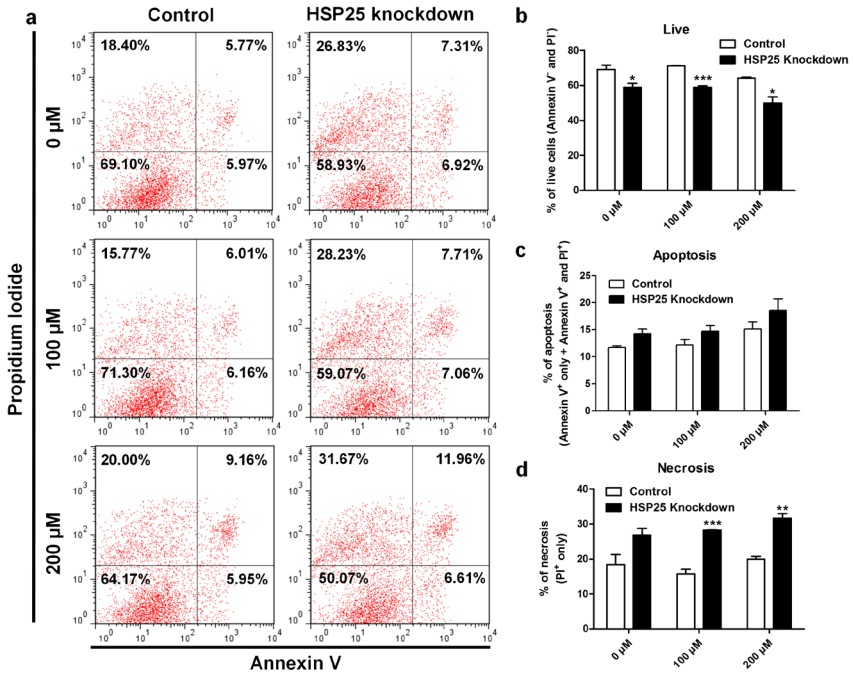


**Figure 7-7**. **Anti-necrotic effect of HSP25 in $H_2O_2$-treated chicken blastoderm cells.** (A) Annexin V/PI analysis by flow cytometry after *HSP25* knockdown followed by $H_2O_2$ treatment (0, 100, or 200 μM). Quantitative analysis of double-negative (B), double-positive (C) and PI-positive only (D) cells, indicating live, apoptotic, and necrotic cells, respectively. Non-complementary sequences in the chicken genome were used as a control. Significant differences between control and *HSP25* knockdown are indicated as ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$. Error bars indicate the SE of triplicate analyses.

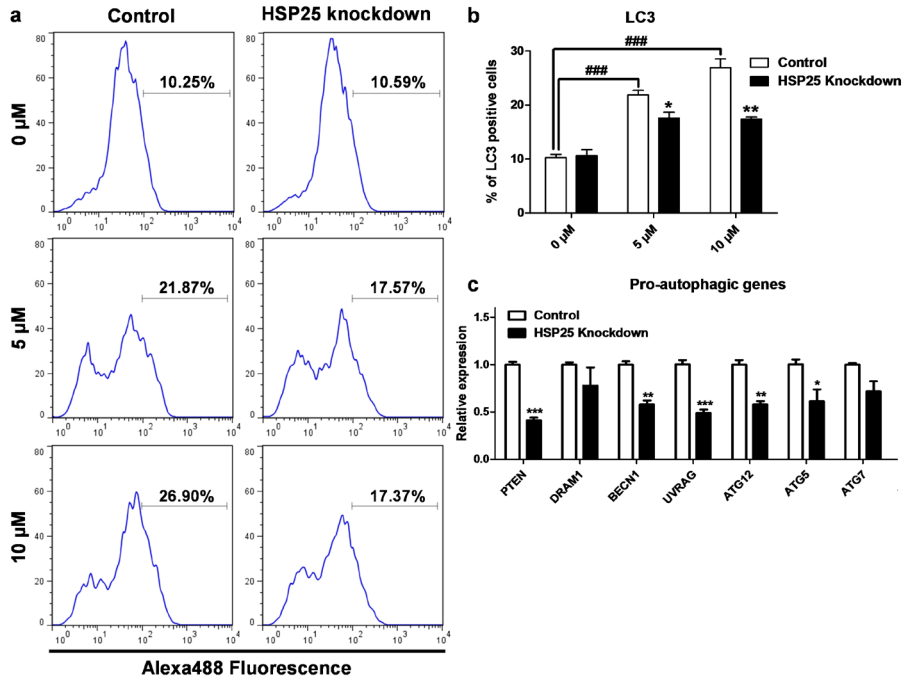


**Figure 7-8**. **Pro-autophagic effect of HSP25 in MG132-treated chicken blastoderm cells.** (A) Flow cytometry analysis of blastoderm cells with LC3 after *HSP25* knockdown followed by MG132 treatment (0, 5, or 10 μM). An Alexa488-conjugated secondary antibody for rabbit IgG was used. (B) Quantitative analysis of LC3-positive cells. (C) Relative expression analysis of pro-autophagic genes after HSP25 knockdown followed by 10 μM MG132 treatment. Non-complementary sequences in the chicken genome were used as a control. Real-time PCR was conducted in triplicate and normalised to expression of *GAPDH*. ###$P < 0.001$ significance of MG132 treatment (5, or 10 μM) compared to 0 μM. Significant differences between control and *HSP25* knockdown are indicated as ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$. Error bars indicate the SE of triplicate analyses.

**Table 7-1. Primers used for quantitative real-time PCR**

| | | Primer sequences | |
|---|---|---|---|
| Accession No. | Gene | Forward (5'→3') | Reverse (5'→3') |
| NM_204305 (Chicken) | *GAPDH* | GGTGGTGCTAAGCGTGTTAT | ACCTCTGTCATCTCTCCACA |
| XM_015873412 (Quail) | *GAPDH* | GGGAAGTTGTGGAGGGATGG | GGTTGGCACACGGAAAGCCA |
| NM_001198610 (Zebra finch) | *GAPDH* | GGGAAGTTGTGGAGGGATGG | GGTTGGCACACGGAAAGCCA |
| NM_205518 | *ACTB* | AGGAGATCACAGCCCTGGCA | CAATGGAGGGTCCGGATTCA |
| NM_001010842 (Chicken) | *HSP25* | AGTTCATGAGCAGCTTCGAG | GTGTTCTGCGTCTCCTTCTG |
| XM_015885650 (Quail) | *HSP25* | AGTTCATGAGCAGCTTCGAG | GTGTTCTGCGTCTCCTTCTG |
| NM_001245736 (Zebra finch) | *HSP25* | AGTTCATGAGCAGCTTCGAG | CTTGTAGGAGAAGGAGCCCT |
| XM_003642832 (Chicken) | *HSP30CL* | GGCAGTTCATGAGCAGCTTC | ACTTGTAGGAGAAGGAGCCC |
| XM_015885649 (Quail) | *HSP30CL* | TGGAGATGGAGAGAGCTCGG | ACTTGTAGGAGAAGGAGCCC |
| XM_004174664 (Zebra finch) | *HSP30CL* | TGGAGATGGAGAGAGCTCGG | ACTTGTAGGAGAAGGAGCCC |
| NM_205290 | *HSPB1* | CGTGCCCTTCACCTTCCTCA | TGGGCAGCAGACGGAAGTATC |
| NM_001001527 | *HSPB2* | GGCCTCGGATCAACAAGCAG | TCTCGGGAGATGAAGCCGTG |
| XM_001231557 | *HSPB3* | CTGCCGAGGAGCTGGAAGAA | CTGGGCGAAACTGCACAACA |
| NM_001030797 | *HSPB4* | TTACCATCCAGCACCCCTGG | GAAATGCCCGACTCCAGCAC |
| NM_205176 | *HSPB5 (CRYAB)* | CCAGGGCCATTCAAGAGCAC | CGTGGCTTCCTTGAGCCAAC |
| XM_427836 | *HSPB7* | CCCGGATGCAAACAAGAAGC | TTGGTGACCTTCCAGGTGGG |
| XM_415280 | *HSPB8* | GCGTCAACGTGCACAGCTTC | AGGCGAAGACAGTGATGGGG |
| XM_418368 | *HSPB10* | GCTTTTTGGAGGACCACGAG | GTTTTCTCCGTACACTGGCC |
| NM_001277613 | *HSPB11* | TGGCTTTCCCAAGAGTGTGA | TCACCCTGTGCACTGATACA |
| NM_001146142 | NANOG | AACTCTGCGGGGCTGTCTTG | AAAAGTGGGGCGGTGAGATG |
| NM_001110178 | *POUV* | TGAAGGGAACGCTGGAGAGC | ATGTCACTGGGATGGGCAGAC |

| | | | |
|---|---|---|---|
| NM_205188 | *SOX2* | CCCCTTTCATACCCCTCCCT | CGGGCTGTTCTTCTGGTTGTT |
| NM_204700 | *CRIPTO* | CGACACAGCGCAAAAGCAAT | AGCCCTTCATCACCCAGTCG |
| NM_205339 | *BCL2* | CTTTATCCTCCTGCCCCTCG | TTCTTCCGCTTCGTCAGCAA |
| NM_001025304 | *BCL2L1* | TTCAGCGACCTCACCTCCCA | GCCCCCAGTTCACACCATCA |
| NM_001031215 | *CAT* | TCAGAAGCCAGATGCCTTGA | CCTTTCGCATGCACAACTCT |
| NM_001277853 | *GPX1* | CATGTTCGAGAAGTGCGAGG | AAGTTCCAGGAGACGTCGTT |
| NM_001277854 | *GPX2* | CTTCTACGACCTCAGTGCCA | TGAGGCAGTTGAGGATCTCC |
| NM_001163232 | *GPX3* | GAACTCCGAGATCCTCCCTG | GGCTCCCAGAAGAGGTTCTT |
| XM_003642871 | *GPX4* | GAGGTGGCGGGATGGGTT | GTACTGCTCCAGGGACACAT |
| NM_001163245 | *GPX7* | AAGCCTTCCTCTTCCTTCCC | GACAACTAACGACACCGAGC |
| XM_423834 | *GPX8* | GGTTGTAAACGTGGCGAGTT | TTCCCTTTGGCAAAAGCCTC |
| XM_421555 | *PTEN* | CCGTTGGTCAACTTTGCTCA | CATGTCCCCAGGTAACCACT |
| XM_416328 | *DRAM1* | GTGGCCTTTGGTTTCGTCTT | TGCTCCTCTGCAATTCTCCA |
| NM_001006332 | *BECN1* | GCTGAAGCTCGATACCTCCT | ATTCTGGCTGGTGGGATGAA |
| NM_001030839 | *UVRAG* | AAGTACTTGGGGCAGCAGAT | AGCTCTGTGAAGCCTAAGCA |
| XM_003643073 | *ATG12* | AGTTGTCTGCTTGCATGGTG | ATCGGTTCCAGTGCCACTTA |
| NM_001006409 | *ATG5* | GGGCATGTTAACTGCTTGCT | CGGTGAAGAAAAGCAGCAGT |
| NM_001030592 | *ATG7* | ATCAAGAGACACAGGCTGCT | TCATGGCAGTCAGAATGGGT |
| NM_001031461 | *MAP1LC3B* | TGTACGAGAGCGAGAAGGAC | GCGTAGGTGACATCCGTTTC |

**Table 7-2. List of *HSP25*-specific siRNA sequences for knockdown experiments**

| Accession No. | Gene | Sequences | Start position |
|---|---|---|---|
| NM_001010842 | *HSP25* | S 5'-GAGAAGGCUCGGCAGUUCA UU-3'<br>AS 5'-UGAACUGCCGAGCCUUCUC UU-3' | 296 |
| | | S 5'-CAGAAGGAGACGCAGAACA UU-3'<br>AS 5'-UGUUCUGCGUCUCCUUCUG UU-3' | 497 |
| N/A | Non-specific siRNA[#] | S 5'-CCUACGCCACCAAUUUCGU-3'<br>AS 5'-GGAUGCGGUGGUUAAAGCA -3' | N/A |

# The non-specific siRNA has no complementary sequence in the chicken genome and was used as a control for gene silencing.

## 4. Discussion

sHSPs are ATP-independent, contain an α-crystallin domain, and prevent the irreversible denaturation of other proteins. sHSPs are known to be responsible for the transfer of other proteins to the ATP-dependent chaperones or to the protein degradation machinery, such as proteasomes or autophagosomes (Mymrikov et al., 2011). In many organisms, sHSPs are accumulated for stress tolerance during diapause and other dormant states in which the ATP concentration is low or limited (Rinehart et al., 2007, MacRae, 2010). Among avian species, newly oviposited EGK stage X chicken blastoderms can endure in a dormant state known as cold torpor for 3-4 weeks (Bloom et al., 1998, Fasenko, 2007). Thus, we hypothesised that chicken blastoderm may show specific sHSP protein expression and function.

First, we examined the expression of 11 *sHSPs*, including the *HSP30* subfamily, in chicken EGK stage X blastoderm cells. Among them, *HSP25* and *HSP30CL* of the HSP30 subfamily were expressed specifically in the chicken blastoderm and *HSP25* showed higher expression than *HSP30CL*. The HSP30 subfamily is absent from mammals and may be restricted to oviparous animals, such as frogs, fish, and avians (Krone et al., 1992, Norris et al., 1997, Katoh et al., 2004). Thus, we identified two sHSPs which may be candidates for regulating dormancy in the chicken blastoderm.

Next, we analyzed the expression patterns of *HSP25* and *HSP30CL* during intrauterine and early development. *HSP25* mRNA was upregulated exclusively after EGK.VIII, shortly before oviposition, whereas *HSP30CL* mRNA was induced after EGK.III, subsequently decreased at EGK.VI, and upregulated again after EGK.VIII. These findings indicated that *HSP25* and *HSP30CL* transcripts were not maternally inherited for embryonic

development and stress tolerance. Additionally, developmentally constitutive *HSP30* homologues in *Xenopus* were first detected in early and mid-tailbud embryos, after gastrulation, but not soon after ZGA, which in *Xenopus* occurs at the 128- to 256-cell stage(Lang et al., 1999). It was recently discovered that ZGA starts between EGK stages II and III in chicken (Nagai et al., 2015, Lee et al., 2016). *HSP25* and *HSP30CL* transcripts were also not induced immediately after ZGA, at EGK III, but earlier than the homologues in *Xenopus*. Accordingly, *HSP25* expression may be related to gaining embryonic tolerance against environmental stress; however, the expression pattern of *HSP30CL* may indicate effects at both the cleavage stage, between EGK.III and VI, and later. Additionally, *HSP25* mRNA was decreased significantly, whereas *HSP30CL* mRNA was upregulated significantly when embryonic development was reinitiated in the 4-h-incubated embryo versus the stage X blastoderm. Species-specific *sHSP* gene expression was increased and expressed exclusively in diapause and dormancy states, such as the dauer stage of *C. elegans* and cysts of *A. franciscana*, but was not detected when development continued (Ludewig et al., 2004, MacRae, 2016). Thus, *HSP25* gene expression was reduced in developmental initiation, but the *HSP30CL* gene was induced, more similar to the gastrulation stages in *Xenopus* (Lang et al., 1999).

Next, we examined *HSP25* and *HSP30CL* expression during egg storage, a dormant state in which energy production was limited. Even in this dormant state under cold storage of eggs for 14 days, *HSP25* expression was upregulated significantly by approximately four-fold after 5 days of storage versus unstored blastoderm, and this level was maintained through 14 days of storage. However, the level of *HSP30CL* expression after storage was only about two-fold higher (not significant). As shown in Figure 7-1A, *HSP25*

transcript levels were the highest of the *sHSPs* in stage X blastoderms. Collectively, these results indicated that, during early development in chicken, *HSP25* was most abundant, and may be specific for embryonic programming of self-defence against future stresses after oviposition.

To identify *HSP25* homologues in vertebrates, we generated multiple alignments and performed phylogenetic tree analysis of sHSPs containing α-crystallin domains HSPB9-like in chicken, Japanese quail, collared flycatcher, Tibetan ground-tit, zebra finch, rock pigeon, frog, salmon, human, and mouse. According to the multiple alignment and phylogenetic trees, the α-crystallin domain in sHSPs was highly conserved, but there was also distinct variations in the N-terminal domain among species. In particular, the N-termini of avian sHSPs differ significantly from those of human, mouse, salmon, and frog sHSPs. Thus, HSP25 and HSP30CL in avian species are quite different to those of other vertebrates.

Although the structure and organisation of the N-terminus in sHSPs is not yet well defined, it is known that the N-terminal domains determine phosphorylation sites, and dimer and oligomer organization (Basha et al., 2012, Mymrikov et al., 2011). Thus, the N-terminal similarity of avian sHSPs could indicate analogous function and substrates. Moreover, based on phylogenetic trees, they can be divided into two groups: one represented by chicken HSP25 and the other by chicken HSP30CL. The similarity between HSP25s versus other sHSPs also showed the same aspect of branching. Thus, avian HSP25 and HSP30CL are distinct sHSPs and may function differently. Furthermore, because of the specific expression of quail and zebra finch *HSP25* and *HSP30CL* in blastoderms at oviposition, and the increased expression of *HSP25* after 7 days of cold storage in both species, which are

similar to that in chicken, HSP25 and HSP30CL homologues may be expressed and act in the same way throughout avian species.

Chicken HSP25, induced by various stresses, has been studied in somatic fibroblast cells (Katoh et al., 2004) and adult tissues (Kawazoe et al., 1999, Wang et al., 2013, Luo et al., 2014, Wang et al., 2015). Chicken HSP25 was first isolated by Kawazoe et al. Subsequently, Katoh et al. observed the accumulation of HSP25 in the aggresomes of somatic fibroblasts. In terms of the genomic structure of chicken *HSP25*, the lack of introns may facilitate rapid expression without disturbance by stressors that could interfere with RNA splicing (Sonna et al., 2002). Also, *HSP25* was the most significant *sHSP* expressed in testes, brain, liver, and egg muscle of egg-laying and broiler adult chickens in response to acute heat stress exposure (Wang et al., 2013, Luo et al., 2014, Wang et al., 2015). Thus, among the *sHSPs* in chicken, *HSP25* could be the first line of the cellular defense against environmental stresses.

Although chicken HSP25 was discovered through accumulation and inclusion formation in chicken somatic cells (Katoh et al., 2004) and adult tissues (Kawazoe et al., 1999, Wang et al., 2013, Luo et al., 2014, Wang et al., 2015), to date, loss of function studies have not been reported. Thus, we next investigated the effects of siRNA-mediated *HSP25* knockdown in chicken blastoderm cells *in vitro* using siRNA-296. Previous studies demonstrated stemness was composed of stress defense, as well as pluripotency in mouse embryonic stem cells. *Hspb1*, a unique signature for stress tolerance, was specific to mouse embryonic stem cells and was reduced along with pluripotency markers by loss of stemness during embryoid bodies formation (Saretzki et al., 2004, Saretzki et al., 2008). Therefore, we examined

pluripotency-related genes, including *NANOG*, *POUV*, *SOX2*, and *CRIPTO* (Lee et al., 2010), in chicken blastoderm cells harboring stemness after *HSP25* knockdown. *In vitro* knockdown of *HSP25* in blastoderm cells caused decreased expression of all pluripotency-related genes, suggesting that *HSP25* seems to have a positive correlation with stem cell property in chickens, like *Hspb1* in mouse embryonic stem cells.

sHSPs are known to be involved in cell survival mechanisms, such as anti-apoptosis (Bruey et al., 2000, Charette et al., 2000, Pandey et al., 2000, Paul et al., 2002, Chauhan et al., 2003, Mao et al., 2004, Havasi et al., 2008), anti-oxidative stress (Mehlen et al., 1996, Preville et al., 1999, Escobedo et al., 2004, Arrigo, 2007), and autophagy (Carra et al., 2008, Carra et al., 2009, Arndt et al., 2010). Cellular homeostasis and integrity should be maintained appropriately for future embryonic development. Nevertheless, the role of chicken *HSP25* and dormancy-specific sHSPs related to biological processes in cell protection have remained largely unknown. We investigated blastoderm cells after *HSP25* knockdown followed by mitomycin C treatment to induce apoptosis. Annexin V/PI analysis showed that suppression of *HSP25* decreased live cells significantly and increased apoptotic cells versus the control. Also, based on qRT-PCR analysis, anti-apoptotic genes were downregulated in knockdown blastoderm cells. These results suggest that chicken *HSP25* in blastoderm cells is important for regulating apoptosis against stress. Additionally, we discovered an anti-oxidant effect of *HSP25*. DCFDA and qRT-PCR analysis indicated that knockdown of *HSP25* followed by $H_2O_2$ treatment increased ROS production significantly and downregulated some anti-oxidant genes in blastoderm cells versus the control. Furthermore, we measured live and dead cells using annexin V/PI flow cytometry. Due to increased ROS production after *HSP25* knockdown, live cells were decreased

significantly, but not apoptotic cells. Rather, necrotic cells were increased, consistent with a previous study indicating ROS-induce necrosis (Samali et al., 1999). Finally, to assess the pro-autophagic function of chicken *HSP25*, we examined autophagy stimulation with the autophagy marker LC3 after inhibition of *HSP25*. Quantification using flow cytometry indicated that autophagic activation was suppressed significantly with *HSP25* knockdown versus the control. Additionally, the relative expression of pro-autophagic genes was downregulated by *HSP25* knockdown. Consequently, chicken HSP25 may be involved in the regulation of processes including apoptosis, anti-oxidative stress, and autophagy for cellular integrity in blastoderm cells.

In conclusion, we found that *HSP25* may be associated with stress-tolerant characteristics in blastoderm cells in chicken. Also, homologues of chicken HSP25 were present in and conserved among avian species. Furthermore, we performed chicken *HSP25* knockdown experiments in blastoderm cells to examine functions associated with apoptosis, anti-oxidative stress, and autophagy. Finally, we demonstrated avian *HSP25* is the first line of cellular defense against environmental stresses and is important in protecting future embryonic cells in the avian blastoderm in cold torpor, as a dormancy state.

# CHAPTER 8

# GENERAL DISCUSSION

In the first study, a total 137 pre-oviposited embryos from oocyte to EGK.X were collected in chicken, a representative model-organism for avian. To investigate the temporal regulation of gene expression, which is important to understand early embryonic development, the whole-transcriptome RNA sequencing of the collected embryos were performed. In addition, transcriptome data for human and mouse early embryogenesis stages (Oocyte to morula), were employed for species comparison among amniotes. As a result, the concept of first- and second- wave of zygotic genome activation (ZGA) was firstly established in chicken early embryo. Functionally, distinct developmental signaling pathways between cleavage and area pellucida formation period were observed and we specified those cell signaling, such as Notch, MAPK, Wnt, and TGF-beta signaling, associated with developmental aspects on aves. In addition to avian developmental programs, species comparison from the evolutionary perspective with transcriptome data for human and mouse early embryogenesis stages (oocyte to morula) highlighted that the chicken-specific features of signaling pathways, as well as the gradually analogous cellular functions such as cell cycle, transcription, and translation among amniotes through ZGA.

Based on the generated transcriptomic information of chicken early embryos, we performed the co-expression analysis based on semi-supervised clustering method to examine the molecular transitions of intrauterine development. As a result, two waves of ZGA-mediated maternal-to-zygotic transition (MZT) were observed across the early embryonic stages and they were associated with transcriptional and translational dynamics. Furthermore, the definite transitions were observed according to the distinct developmental characteristics between cleavage and area pellucida formation period in the functional analysis. Finally, epigenetic modification and the evolutionarily

conserved miRNA expression suggest that the certain MZT proceeds from stage EGK.VIII in early chicken development. Consequently, together with morphological perspectives indicating mid-blastula transition (MBT), we firstly demonstrated the transcriptomic transitions of early developmental stages and the establishment of embryonic identify during MZT in EGK.VIII stage of chicken.

Next, our transcriptomic analysis of early embryos profiled the DNA-dependent transcription factors (TFs) in chicken. As a result, the expression of TFs by the two waves ZGA before the beginning of the distinct developmental periods seemed to be ready to govern cleavage and area pellucida formation during intrauterine development, respectively. This study provides the transcriptional regulation and DNA binding TFs during early embryonic development in avian species. Also, our expression profiling of TFs and the limitation of iPSC-like cells in birds provide the new insight into the reprogramming to totipotency and pluripotency in avian species.

To look more deeply into ZGA in chicken, our genome-wide study of primary transcripts, traced by intronic mapped reads, clarified that the two definite waves of transcriptional activation including 1st wave after fertilization. Furthermore, our additional multi-omics approaches revealed an interesting result that all transcripts in the zygote stage come from the maternal genome activation (MGA) exclusively, even with first-wave of transcriptional activation. We also speculate that this evolved by necessity in animals with physiological polyspermy. To inhibit disproportionate genome contribution or genetic instability due to inattentive transcription of paternal alleles, there may be the feasible mechanism of allelic histone modification in avian zygotes activating only maternal alleles after fertilisation, which need to

be revealed.

Finally, we discovered the avian-specific small heat shock protein (sHSP), HSP25, in blastoderm dormancy. Among 11 sHSPs of chicken, *HSP25* was expressed especially in the blastoderm and was highly upregulated in dormant state. Next, we found homologues of HSP25 were conserved in other avian species. The blastoderm of Japanese quail and zebra finch showed similar expression pattern of HSP25 to chicken. Also, loss of function studies demonstrated that chicken HSP25 is associated with pluripotency, anti-apoptotic, anti-oxidant, and pro-autophagic effects in blastodermal cells. In conclusion, our results suggest that avian HSP25 is important in the defence against environmental stress and the protection of future embryonic cells in the avian blastoderm. This study provide that the constitutive sHSP expression is related with blastoderm dormancy in avian species for the first time.

# REFERENCES

AANES, H., WINATA, C. L., LIN, C. H., CHEN, J. P., SRINIVASAN, K. G., LEE, S. G., LIM, A. Y., HAJAN, H. S., COLLAS, P., BOURQUE, G., GONG, Z., KORZH, V., ALESTROM, P. & MATHAVAN, S. 2011. Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition. Genome Res, 21, 1328-38.

ABE, K., YAMAMOTO, R., FRANKE, V., CAO, M., SUZUKI, Y., SUZUKI, M. G., VLAHOVICEK, K., SVOBODA, P., SCHULTZ, R. M. & AOKI, F. 2015. The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3' processing. EMBO J, 34, 1523-37.

ACLOQUE, H., RISSON, V., BIROT, A. M., KUNITA, R., PAIN, B. & SAMARUT, J. 2001. Identification of a new gene family specifically expressed in chicken embryonic stem cells and early embryo. Mech Dev, 103, 79-91.

AKKERS, R. C., VAN HEERINGEN, S. J., JACOBI, U. G., JANSSEN-MEGENS, E. M., FRANCOIJS, K. J., STUNNENBERG, H. G. & VEENSTRA, G. J. 2009. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos. Dev Cell, 17, 425-34.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics, 31, 166-169.

ANDERSON, S. N., JOHNSON, C. S., CHESNUT, J., JONES, D. S.,

KHANDAY, I., WOODHOUSE, M., LI, C., CONRAD, L. J., RUSSELL, S. D. & SUNDARESAN, V. 2017. The Zygotic Transition Is Initiated in Unicellular Plant Zygotes with Asymmetric Activation of Parental Genomes. Dev Cell, 43, 349-358 e4.

ANDREWS, S. 2010. FastQC: A quality control tool for high throughput sequence data. Reference Source.

ANDRIES, L., VAKAET, L. & VANROELEN, C. 1983. The subgerminal yolk surface and its relationship with the inner germ wall edge of the stages X to XIV chick and quail embryo. A SEM study. Anat Embryol (Berl), 166, 453-62.

AOKI, F., WORRAD, D. M. & SCHULTZ, R. M. 1997. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. Dev Biol, 181, 296-307.

AOSHIMA, K., INOUE, E., SAWA, H. & OKADA, Y. 2015. Paternal H3K4 methylation is required for minor zygotic gene activation and early mouse embryonic development. EMBO Rep, 16, 803-12.

ARNDT, V., DICK, N., TAWO, R., DREISEIDLER, M., WENZEL, D., HESSE, M., FURST, D. O., SAFTIG, P., SAINT, R., FLEISCHMANN, B. K., HOCH, M. & HOHFELD, J. 2010. Chaperone-assisted selective autophagy is essential for muscle maintenance. Curr Biol, 20, 143-8.

ARRIGO, A. P. 2007. The cellular "networking" of mammalian Hsp27 and its functions in the control of protein folding, redox state and apoptosis. Adv Exp Med Biol, 594, 14-26.

BASHA, E., O'NEILL, H. & VIERLING, E. 2012. Small heat shock proteins and alpha-crystallins: dynamic proteins with flexible functions. Trends Biochem Sci, 37, 106-17.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), 289-300.

BIRKHEAD, T. R., SHELDON, B. C. & FLETCHER, F. 1994. A comparative study of sperm-egg interactions in birds. J Reprod Fertil, 101, 353-61.

BLOOM, S. E., MUSCARELLA, D. E., LEE, M. Y. & RACHLINSKI, M. 1998. Cell death in the avian blastoderm: resistance to stress-induced apoptosis and expression of anti-apoptotic genes. Cell Death Differ, 5, 529-38.

BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, btu170.

BOUNIOL, C., NGUYEN, E. & DEBEY, P. 1995. Endogenous transcription occurs at the 1-cell stage in the mouse embryo. Exp Cell Res, 218, 57-62.

BRAUDE, P., BOLTON, V. & MOORE, S. 1988. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. Nature, 332, 459-61.

BRAUDE, P., PELHAM, H., FLACH, G. & LOBATTO, R. 1979. Post-transcriptional control in the early mouse embryo. Nature, 282, 102-5.

BRUEY, J. M., DUCASSE, C., BONNIAUD, P., RAVAGNAN, L., SUSIN, S. A., DIAZ-LATOUD, C., GURBUXANI, S., ARRIGO, A. P., KROEMER, G., SOLARY, E. & GARRIDO, C. 2000. Hsp27 negatively regulates cell death by interacting with cytochrome c. Nat Cell Biol, 2, 645-52.

BULUT-KARSLIOGLU, A., BIECHELE, S., JIN, H., MACRAE, T. A., HEJNA, M., GERTSENSTEIN, M., SONG, J. S. & RAMALHO-SANTOS, M. 2016. Inhibition of mTOR induces a paused pluripotent state. Nature, 540, 119-123.

CADIGAN, K. M. & WATERMAN, M. L. 2012. TCF/LEFs and Wnt signaling in the nucleus. Cold Spring Harb Perspect Biol, 4.

CALLEBAUT, M. 2005. Origin, fate, and function of the components of the avian germ disc region and early blastoderm: role of ooplasmic determinants. Dev Dyn, 233, 1194-216.

CANTONE, I. & FISHER, A. G. 2013. Epigenetic programming and reprogramming during development. Nat Struct Mol Biol, 20, 282-9.

CARRA, S., BRUNSTING, J. F., LAMBERT, H., LANDRY, J. & KAMPINGA, H. H. 2009. HspB8 participates in protein quality control by a non-chaperone-like mechanism that requires eIF2{alpha} phosphorylation. J Biol Chem, 284, 5523-32.

CARRA, S., SEGUIN, S. J., LAMBERT, H. & LANDRY, J. 2008. HspB8 chaperone activity toward poly(Q)-containing proteins depends on its association with Bag3, a stimulator of macroautophagy. J Biol Chem, 283, 1437-44.

CASTANON, I., ABRAMI, L., HOLTZER, L., HEISENBERG, C. P., VAN DER GOOT, F. G. & GONZALEZ-GAITAN, M. 2013. Anthrax toxin receptor 2a controls mitotic spindle positioning. Nat Cell Biol, 15, 28-39.

CHA, S. W., TADJUIDJE, E., TAO, Q., WYLIE, C. & HEASMAN, J. 2008. Wnt5a and Wnt11 interact in a maternal Dkk1-regulated fashion to activate both canonical and non-canonical signaling in Xenopus axis formation. *Development,* 135**,** 3719-29.

CHARETTE, S. J., LAVOIE, J. N., LAMBERT, H. & LANDRY, J. 2000. Inhibition of Daxx-mediated apoptosis by heat shock protein 27. Mol Cell Biol, 20, 7602-12.

CHAUHAN, D., LI, G., HIDESHIMA, T., PODAR, K., MITSIADES, C., MITSIADES, N., CATLEY, L., TAI, Y. T., HAYASHI, T., SHRINGARPURE, R., BURGER, R., MUNSHI, N., OHTAKE, Y., SAXENA, S. & ANDERSON, K. C. 2003. Hsp27 inhibits release of mitochondrial protein Smac in multiple myeloma cells and confers dexamethasone resistance. Blood, 102, 3379-86.

CHEN, J., STRIEDER, N., KROHN, N. G., CYPRYS, P., SPRUNCK, S., ENGELMANN, J. C. & DRESSELHAUS, T. 2017. Zygotic Genome Activation Occurs Shortly after Fertilization in Maize. Plant Cell, 29, 2106-2125.

CHOI, H. W., KIM, J. S., CHOI, S., JU HONG, Y., BYUN, S. J., SEO, H. G. & DO, J. T. 2016. Mitochondrial Remodeling in Chicken Induced Pluripotent Stem-Like Cells. Stem Cells Dev, 25, 472-6.

CHOI, J. W., LEE, E. Y., SHIN, J. H., ZHENG, Y., CHO, B. W., KIM, J. K.,

KIM, H. & HAN, J. Y. 2007. Identification of breed-specific DNA polymorphisms for a simple and unambiguous screening system in germline chimeric chickens. J Exp Zool A Ecol Genet Physiol, 307, 241-8.

CINGOLANI, P., PATEL, V. M., COON, M., NGUYEN, T., LAND, S. J., RUDEN, D. M. & LU, X. 2012. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. Frontiers in genetics, 3.

DAHL, J. A., JUNG, I., AANES, H., GREGGAINS, G. D., MANAF, A., LERDRUP, M., LI, G., KUAN, S., LI, B., LEE, A. Y., PREISSL, S., JERMSTAD, I., HAUGEN, M. H., SUGANTHAN, R., BJORAS, M., HANSEN, K., DALEN, K. T., FEDORCSAK, P., REN, B. & KLUNGLAND, A. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature, 537, 548-552.

DENICOL, A. C., BLOCK, J., KELLEY, D. E., POHLER, K. G., DOBBS, K. B., MORTENSEN, C. J., ORTEGA, M. S. & HANSEN, P. J. 2014. The WNT signaling antagonist Dickkopf-1 directs lineage commitment and promotes survival of the preimplantation embryo. *FASEB J,* 28**,** 3975-86.

DENICOL, A. C., DOBBS, K. B., MCLEAN, K. M., CARAMBULA, S. F., LOUREIRO, B. & HANSEN, P. J. 2013. Canonical WNT signaling regulates development of bovine embryos to the blastocyst stage. *Sci Rep,* 3**,** 1266.

EDGAR, B. A. & DATAR, S. A. 1996. Zygotic degradation of two maternal Cdc25 mRNAs terminates Drosophila's early cell cycle program. Genes Dev, 10, 1966-77.

ELIS, S., BATELLIER, F., COUTY, I., BALZERGUE, S., MARTIN-MAGNIETTE, M. L., MONGET, P., BLESBOIS, E. & GOVOROUN, M. S. 2008. Search for the genes involved in oocyte maturation and early embryo development in the hen. BMC Genomics, 9, 110.

ESCOBEDO, J., PUCCI, A. M. & KOH, T. J. 2004. HSP25 protects skeletal muscle cells against oxidative stress. Free Radic Biol Med, 37, 1455-62.

EYAL-GILADI, H. & KOCHAV, S. 1976. From cleavage to primitive streak formation: a complementary normal table and a new look at the first stages of the development of the chick. I. General morphology. Dev Biol, 49, 321-37.

FABIAN, B. & EYAL-GILADI, H. 1981. A SEM study of cell shedding during the formation of the area pellucida in the chick embryo. J Embryol Exp Morphol, 64, 11-22.

FASENKO, G. M. 2007. Egg storage and the embryo. Poult Sci, 86, 1020-4.

FENELON, J. C., BANERJEE, A. & MURPHY, B. D. 2014. Embryonic diapause: development on hold. Int J Dev Biol, 58, 163-74.

FRESARD, L., LEROUX, S., SERVIN, B., GOURICHON, D., DEHAIS, P., CRISTOBAL, M. S., MARSAUD, N., VIGNOLES, F., BED'HOM, B., COVILLE, J. L., HORMOZDIARI, F., BEAUMONT, C., ZERJAL, T., VIGNAL, A., MORISSON, M., LAGARRIGUE, S. & PITEL, F. 2014. Transcriptome-wide investigation of genomic imprinting in chicken. Nucleic Acids Res, 42, 3768-82.

GARDNER, A. J. & EVANS, J. P. 2006. Mammalian membrane block to

polyspermy: new insights into how mammalian eggs prevent fertilisation by multiple sperm. Reprod Fertil Dev, 18, 53-61.

GEISER, F. 2004. Metabolic rate and body temperature reduction during hibernation and daily torpor. Annu. Rev. Physiol., 66, 239-274.

GILBERT, I., SCANTLAND, S., SYLVESTRE, E. L., GRAVEL, C., LAFLAMME, I., SIRARD, M. A. & ROBERT, C. 2009. The dynamics of gene products fluctuation during bovine pre-hatching development. Mol Reprod Dev, 76, 762-72.

GIRALDEZ, A. J., MISHIMA, Y., RIHEL, J., GROCOCK, R. J., VAN DONGEN, S., INOUE, K., ENRIGHT, A. J. & SCHIER, A. F. 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science, 312, 75-9.

GOMEZ-DE-TRAVECEDO, P., CARAVACA, F. P. & GONZALEZ-REDONDO, P. 2014. Effects of storage temperature and length of the storage period on hatchability and performance of red-legged partridge (Alectoris rufa) eggs. Poult Sci, 93, 747-54.

GRAF, A., KREBS, S., ZAKHARTCHENKO, V., SCHWALB, B., BLUM, H. & WOLF, E. 2014. Fine mapping of genome activation in bovine embryos by RNA sequencing. Proc Natl Acad Sci U S A, 111, 4139-44.

GRAF, T., MCNAGNY, K., BRADY, G. & FRAMPTON, J. 1992. Chicken "erythroid" cells transformed by the Gag-Myb-Ets-encoding E26 leukemia virus are multipotent. Cell, 70, 201-13.

GUSSE, M., GHYSDAEL, J., EVAN, G., SOUSSI, T. & MECHALI, M. 1989. Translocation of a store of maternal cytoplasmic c-myc protein into nuclei during early development. Mol Cell Biol, 9, 5395-403.

HAMBURGER, V. & HAMILTON, H. L. 1951. A series of normal stages in the development of the chick embryo. J Morphol, 88, 49-92.

HAMIDU, J. A., RIEGER, A. M., FASENKO, G. M. & BARREDA, D. R. 2010. Dissociation of chicken blastoderm for examination of apoptosis and necrosis by flow cytometry. Poult Sci, 89, 901-9.

HAN, J. Y. 2009. Germ cells and transgenesis in chickens. Comp Immunol Microbiol Infect Dis, 32, 61-80.

HARVEY, S. A., SEALY, I., KETTLEBOROUGH, R., FENYES, F., WHITE, R., STEMPLE, D. & SMITH, J. C. 2013. Identification of the zebrafish maternal and paternal transcriptomes. Development, 140, 2703-10.

HAVASI, A., LI, Z., WANG, Z., MARTIN, J. L., BOTLA, V., RUCHALSKI, K., SCHWARTZ, J. H. & BORKAN, S. C. 2008. Hsp27 inhibits Bax activation and apoptosis via a phosphatidylinositol 3-kinase-dependent mechanism. J Biol Chem, 283, 12305-13.

HAVASI, A., LI, Z., WANG, Z., MARTIN, J. L., BOTLA, V., RUCHALSKI, K., SCHWARTZ, J. H. & BORKAN, S. C. 2008. Hsp27 inhibits Bax activation and apoptosis via a phosphatidylinositol 3-kinase-dependent mechanism. *J Biol Chem,* 283**,** 12305-13.

HEMMINGS, N. & BIRKHEAD, T. R. 2015. Polyspermy in birds: sperm

numbers and embryo survival. Proc Biol Sci, 282, 20151682.

HEYN, P., KIRCHER, M., DAHL, A., KELSO, J., TOMANCAK, P., KALINKA, A. T. & NEUGEBAUER, K. M. 2014. The earliest transcribed zygotic genes are short, newly evolved, and different across species. Cell Rep, 6, 285-92.

HIKASA, H. & SOKOL, S. Y. 2013. Wnt signaling in vertebrate axis specification. *Cold Spring Harb Perspect Biol,* 5**,** a007955.

HSU, A. L., MURPHY, C. T. & KENYON, C. 2003. Regulation of aging and age-related disease by DAF-16 and heat-shock factor. Science, 300, 1142-5.

HUANG, D. W., SHERMAN, B. T., TAN, Q., KIR, J., LIU, D., BRYANT, D., GUO, Y., STEPHENS, R., BASELER, M. W. & LANE, H. C. 2007. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic acids research, 35, W169-W175.

HUANG, Y. L., ANVARIAN, Z., DODERLEIN, G., ACEBRON, S. P. & NIEHRS, C. 2015. Maternal Wnt/STOP signaling promotes cell division during early Xenopus embryogenesis. Proc Natl Acad Sci U S A, 112, 5732-7.

HUNTER, S., APWEILER, R., ATTWOOD, T. K., BAIROCH, A., BATEMAN, A., BINNS, D., BORK, P., DAS, U., DAUGHERTY, L. & DUQUENNE, L. 2009. InterPro: the integrative protein signature database. Nucleic acids research, 37, D211-D215.

HWANG, Y. S., SEO, M., BANG, S., KIM, H. & HAN, J. Y. 2018.

Transcriptional and translational dynamics during maternal-to-zygotic transition in early chicken development. FASEB J, 32.

IHARA, M., TSENG, H. & SCHULTZ, R. M. 2011. Expression of variant ribosomal RNA genes in mouse oocytes and preimplantation embryos. Biol Reprod, 84, 944-6.

ISHIUCHI, T., ENRIQUEZ-GASCA, R., MIZUTANI, E., BOSKOVIC, A., ZIEGLER-BIRLING, C., RODRIGUEZ-TERRONES, D., WAKAYAMA, T., VAQUERIZAS, J. M. & TORRES-PADILLA, M. E. 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. Nat Struct Mol Biol, 22, 662-71.

IWAO, Y. 2012. Egg activation in physiological polyspermy. Reproduction, 144, 11-22.

JAMES, D., LEVINE, A. J., BESSER, D. & HEMMATI-BRIVANLOU, A. 2005. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development,* 132**,** 1273-82.

JARVIS, E. D., MIRARAB, S., ABERER, A. J., LI, B., HOUDE, P., LI, C., HO, S. Y., FAIRCLOTH, B. C., NABHOLZ, B., HOWARD, J. T., SUH, A., WEBER, C. C., DA FONSECA, R. R., LI, J., ZHANG, F., LI, H., ZHOU, L., NARULA, N., LIU, L., GANAPATHY, G., BOUSSAU, B., BAYZID, M. S., ZAVIDOVYCH, V., SUBRAMANIAN, S., GABALDON, T., CAPELLA-GUTIERREZ, S., HUERTA-CEPAS, J., REKEPALLI, B., MUNCH, K., SCHIERUP, M., LINDOW, B., WARREN, W. C., RAY, D., GREEN, R. E., BRUFORD, M. W., ZHAN, X., DIXON, A., LI, S., LI, N., HUANG, Y., DERRYBERRY, E. P., BERTELSEN, M. F., SHELDON, F. H.,

BRUMFIELD, R. T., MELLO, C. V., LOVELL, P. V., WIRTHLIN, M., SCHNEIDER, M. P., PROSDOCIMI, F., SAMANIEGO, J. A., VARGAS VELAZQUEZ, A. M., ALFARO-NUNEZ, A., CAMPOS, P. F., PETERSEN, B., SICHERITZ-PONTEN, T., PAS, A., BAILEY, T., SCOFIELD, P., BUNCE, M., LAMBERT, D. M., ZHOU, Q., PERELMAN, P., DRISKELL, A. C., SHAPIRO, B., XIONG, Z., ZENG, Y., LIU, S., LI, Z., LIU, B., WU, K., XIAO, J., YINQI, X., ZHENG, Q., ZHANG, Y., YANG, H., WANG, J., SMEDS, L., RHEINDT, F. E., BRAUN, M., FJELDSA, J., ORLANDO, L., BARKER, F. K., JONSSON, K. A., JOHNSON, W., KOEPFLI, K. P., O'BRIEN, S., HAUSSLER, D., RYDER, O. A., RAHBEK, C., WILLERSLEV, E., GRAVES, G. R., GLENN, T. C., MCCORMACK, J., BURT, D., ELLEGREN, H., ALSTROM, P., EDWARDS, S. V., STAMATAKIS, A., MINDELL, D. P., CRACRAFT, J., et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science, 346, 1320-31.

KATOH, Y., FUJIMOTO, M., NAKAMURA, K., INOUYE, S., SUGAHARA, K., IZU, H. & NAKAI, A. 2004. Hsp25, a member of the Hsp30 family, promotes inclusion formation in response to stress. FEBS Lett, 565, 28-32.

KAWAZOE, Y., TANABE, M. & NAKAI, A. 1999. Ubiquitous and cell-specific members of the avian small heat shock protein family. FEBS Lett, 455, 271-5.

KELLER, R. 2002. Shaping the vertebrate body plan by polarized embryonic cell movements. Science, 298, 1950-4.

KENDALL, M. G., STUART, A. & ORD, J. 1968. The advanced theory of

statistics, London.

KIM, D., LANGMEAD, B. & SALZBERG, S. L. 2015. HISAT: a fast spliced aligner with low memory requirements. Nature methods, 12, 357-360.

KIM, M., GEUM, D., KHANG, I., PARK, Y. M., KANG, B. M., LEE, K. A. & KIM, K. 2002. Expression pattern of HSP25 in mouse preimplantation embryo: heat shock responses during oocyte maturation. Mol Reprod Dev, 61, 3-13.

KIM, Y. M., PARK, Y. H., LIM, J. M., JUNG, H. & HAN, J. Y. 2017. Technical note: Induction of pluripotent stem cell-like cells from chicken feather follicle cells. J Anim Sci, 95, 3479-3486.

KING, A. M. & MACRAE, T. H. 2012. The small heat shock protein p26 aids development of encysting Artemia embryos, prevents spontaneous diapause termination and protects against stress. PLoS One, 7, e43723.

KING, A. M., TOXOPEUS, J. & MACRAE, T. H. 2013. Functional differentiation of small heat shock proteins in diapause-destined Artemia embryos. FEBS J, 280, 4761-72.

KISIELEWSKA, J., PHILIPOVA, R., HUANG, J. Y. & WHITAKER, M. 2009. MAP kinase dependent cyclinE/cdk2 activity promotes DNA replication in early sea urchin embryos. Dev Biol, 334, 383-94.

KISPERT, A., ORTNER, H., COOKE, J. & HERRMANN, B. G. 1995. The chick Brachyury gene: developmental expression pattern and response to axial induction by localized activin. Dev Biol, 168, 406-15.

KO, M. H., HWANG, Y. S., RIM, J. S., HAN, H. J. & HAN, J. Y. 2017. Avian blastoderm dormancy arrests cells in G2 and suppresses apoptosis. FASEB J, 31, 3240-3250.

KOCHAV, S. & EYAL-GILADI, H. 1971. Bilateral symmetry in chick embryo determination by gravity. Science, 171, 1027-9.

KOCHAV, S., GINSBURG, M. & EYAL-GILADI, H. 1980. From cleavage to primitive streak formation: a complementary normal table and a new look at the first stages of the development of the chick. II. Microscopic anatomy and cell population dynamics. Dev Biol, 79, 296-308.

KRONE, P. H., SNOW, A., ALI, A., PASTERNAK, J. J. & HEIKKILA, J. J. 1992. Comparison of regulatory and structural regions of the Xenopus laevis small heat-shock protein-encoding gene family. Gene, 110, 159-66.

LANG, L., MISKOVIC, D., FERNANDO, P. & HEIKKILA, J. J. 1999. Spatial pattern of constitutive and heat shock-induced expression of the small heat shock protein gene family, Hsp30, in Xenopus laevis tailbud embryos. Dev Genet, 25, 365-74.

LANGFELDER, P. & HORVATH, S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics, 9, 1.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods, 9, 357-359.

LEE, H. C., CHOI, H. J., LEE, H. G., LIM, J. M., ONO, T. & HAN, J. Y. 2016. DAZL Expression Explains Origin and Central Formation of Primordial

Germ Cells in Chickens. Stem Cells Dev, 25, 68-79.

LEE, H. C., CHOI, H. J., PARK, T. S., LEE, S. I., KIM, Y. M., RENGARAJ, D., NAGAI, H., SHENG, G., LIM, J. M. & HAN, J. Y. 2013a. Cleavage events and sperm dynamics in chick intrauterine embryos. PLoS One, 8, e80631.

LEE, M. T., BONNEAU, A. R. & GIRALDEZ, A. J. 2014. Zygotic genome activation during the maternal-to-zygotic transition. Annu Rev Cell Dev Biol, 30, 581-613.

LEE, M. T., BONNEAU, A. R., TAKACS, C. M., BAZZINI, A. A., DIVITO, K. R., FLEMING, E. S. & GIRALDEZ, A. J. 2013b. Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature, 503, 360-4.

LEE, S. I., KIM, J. K., PARK, H. J., JANG, H. J., LEE, H. C., MIN, T., SONG, G. & HAN, J. Y. 2010. Molecular cloning and characterization of the germ cell-related nuclear orphan receptor in chickens. Mol Reprod Dev, 77, 273-84.

LEE, S. I., LEE, B. R., HWANG, Y. S., LEE, H. C., RENGARAJ, D., SONG, G., PARK, T. S. & HAN, J. Y. 2011. MicroRNA-mediated posttranscriptional regulation is required for maintaining undifferentiated properties of blastoderm and primordial germ cells in chickens. Proc Natl Acad Sci U S A, 108, 10426-31.

LEICHSENRING, M., MAES, J., MOSSNER, R., DRIEVER, W. & ONICHTCHOUK, D. 2013. Pou5f1 transcription factor controls zygotic gene

activation in vertebrates. Science, 341, 1005-9.

LETUNIC, I., DOERKS, T. & BORK, P. 2011. SMART 7: recent updates to the protein domain annotation resource. Nucleic acids research, 40, D302-D305.

LI, A., ZHANG, J., ZHOU, Z., WANG, L., LIU, Y. & LIU, Y. 2015. ALDB: a domestic-animal long noncoding RNA database. PloS one, 10, e0124003.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics, 25, 2078-2079.

LINDEMAN, L. C., ANDERSEN, I. S., REINER, A. H., LI, N., AANES, H., OSTRUP, O., WINATA, C., MATHAVAN, S., MULLER, F., ALESTROM, P. & COLLAS, P. 2011. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. Dev Cell, 21, 993-1004.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods, 25, 402-8.

LONGO, F. J. & KUNKLE, M. 1977. Synthesis of RNA by male pronuclei of fertilized sea urchin eggs. J Exp Zool, 201, 431-7.

LOPES, F. L., DESMARAIS, J. A. & MURPHY, B. D. 2004. Embryonic diapause and its regulation. Reproduction, 128, 669-78.

LU, Y., WEST, F. D., JORDAN, B. J., JORDAN, E. T., WEST, R. C., YU, P., HE, Y., BARRIOS, M. A., ZHU, Z., PETITTE, J. N., BECKSTEAD, R. B. & STICE, S. L. 2014. Induced pluripotency in chicken embryonic fibroblast results in a germ cell fate. Stem Cells Dev, 23, 1755-64.

LU, Y., WEST, F. D., JORDAN, B. J., MUMAW, J. L., JORDAN, E. T., GALLEGOS-CARDENAS, A., BECKSTEAD, R. B. & STICE, S. L. 2012. Avian-induced pluripotent stem cells derived using human reprogramming factors. Stem Cells Dev, 21, 394-403.

LUDEWIG, A. H., KOBER-EISERMANN, C., WEITZEL, C., BETHKE, A., NEUBERT, K., GERISCH, B., HUTTER, H. & ANTEBI, A. 2004. A novel nuclear receptor/coregulator complex controls C. elegans lipid metabolism, larval development, and aging. Genes Dev, 18, 2120-33.

LUND, E., LIU, M., HARTLEY, R. S., SHEETS, M. D. & DAHLBERG, J. E. 2009. Deadenylation of maternal mRNAs mediated by miR-427 in Xenopus laevis embryos. RNA, 15, 2351-63.

LUO, Q. B., SONG, X. Y., JI, C. L., ZHANG, X. Q. & ZHANG, D. X. 2014. Exploring the molecular mechanism of acute heat stress exposure in broiler chickens using gene expression profiling. Gene, 546, 200-5.

MACFARLAN, T. S., GIFFORD, W. D., DRISCOLL, S., LETTIERI, K., ROWE, H. M., BONANOMI, D., FIRTH, A., SINGER, O., TRONO, D. & PFAFF, S. L. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature, 487, 57-63.

MACRAE, T. H. 2010. Gene expression, metabolic regulation and stress

tolerance during diapause. Cell Mol Life Sci, 67, 2405-24.

MACRAE, T. H. 2016. Stress tolerance during diapause and quiescence of the brine shrimp, Artemia. Cell Stress Chaperones, 21, 9-18.

MAJUMDER, S., MIRANDA, M. & DEPAMPHILIS, M. L. 1993. Analysis of gene expression in mouse preimplantation embryos demonstrates that the primary role of enhancers is to relieve repression of promoters. EMBO J, 12, 1131-40.

MAK, S. S., ALEV, C., NAGAI, H., WRABEL, A., MATSUOKA, Y., HONDA, A., SHENG, G. & LADHER, R. K. 2015. Characterization of the finch embryo supports evolutionary conservation of the naive stage of development in amniotes. Elife, 4, e07178.

MALLO, M. & ALONSO, C. R. 2013. The regulation of Hox gene expression during animal development. Development, 140, 3951-63.

MAO, L. & SHELDEN, E. A. 2006. Developmentally regulated gene expression of the small heat shock protein Hsp27 in zebrafish embryos. Gene Expr Patterns, 6, 127-33.

MAO, Y. W., LIU, J. P., XIANG, H. & LI, D. W. 2004. Human alphaA- and alphaB-crystallins bind to Bax and Bcl-X(S) to sequester their translocation during staurosporine-induced apoptosis. Cell Death Differ, 11, 512-26.

MCAVEY, B. A., WORTZMAN, G. B., WILLIAMS, C. J. & EVANS, J. P. 2002. Involvement of calcium signaling and the actin cytoskeleton in the membrane block to polyspermy in mouse eggs. Biol Reprod, 67, 1342-52.

MCELWEE, J. J., SCHUSTER, E., BLANC, E., THOMAS, J. H. & GEMS, D. 2004. Shared transcriptional signature in Caenorhabditis elegans Dauer larvae and long-lived daf-2 mutants implicates detoxification system in longevity assurance. J Biol Chem, 279, 44533-43.

MCKENNA, A., HANNA, M., BANKS, E., SIVACHENKO, A., CIBULSKIS, K., KERNYTSKY, A., GARIMELLA, K., ALTSHULER, D., GABRIEL, S. & DALY, M. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research, 20, 1297-1303.

MEHLEN, P., KRETZ-REMY, C., PREVILLE, X. & ARRIGO, A. P. 1996. Human hsp27, Drosophila hsp27 and human alphaB-crystallin expression-mediated increase in glutathione is essential for the protective activity of these proteins against TNFalpha-induced cell death. EMBO J, 15, 2695-706.

METZ, T. & GRAF, T. 1991. v-myb and v-ets transform chicken erythroid cells and cooperate both in trans and in cis to induce distinct differentiation phenotypes. Genes Dev, 5, 369-80.

MEY, A., ACLOQUE, H., LERAT, E., GOUNEL, S., TRIBOLLET, V., BLANC, S., CURTON, D., BIROT, A. M., NIETO, M. A. & SAMARUT, J. 2012. The endogenous retrovirus ENS-1 provides active binding sites for transcription factors in embryonic stem cells that specify extra embryonic tissue. Retrovirology, 9, 21.

MIZUSHIMA, S., HIYAMA, G., SHIBA, K., INABA, K., DOHRA, H., ONO, T., SHIMADA, K. & SASANAMI, T. 2014. The birth of quail chicks after intracytoplasmic sperm injection. Development, 141, 3799-806.

MORROW, G. & TANGUAY, R. M. 2012. Small heat shock protein expression and functions during development. Int J Biochem Cell Biol, 44, 1613-21.

MYMRIKOV, E. V., SEIT-NEBI, A. S. & GUSEV, N. B. 2011. Large potentials of small heat shock proteins. Physiol Rev, 91, 1123-59.

NAGAI, H., SEZAKI, M., KAKIGUCHI, K., NAKAYA, Y., LEE, H. C., LADHER, R., SASANAMI, T., HAN, J. Y., YONEMURA, S. & SHENG, G. 2015. Cellular analysis of cleavage-stage chick embryos reveals hidden conservation in vertebrate early development. Development, 142, 1279-86.

NAKANOH, S., OKAZAKI, K. & AGATA, K. 2013. Inhibition of MEK and GSK3 supports ES cell-like domed colony formation from avian and reptile embryos. Zoolog Sci, 30, 543-52.

NEWPORT, J. & KIRSCHNER, M. 1982. A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. Cell, 30, 687-96.

NIEHRS, C. 2012. The complex world of WNT receptor signalling. *Nat Rev Mol Cell Biol,* 13**,** 767-79.

NORRIS, C. E., BROWN, M. A., HICKEY, E., WEBER, L. A. & HIGHTOWER, L. E. 1997. Low-molecular-weight heat shock proteins in a desert fish (Poeciliopsis lucida): homologs of human Hsp27 and Xenopus Hsp30. Mol Biol Evol, 14, 1050-61.

OHGUSHI, M. & SASAI, Y. 2011. Lonely death dance of human pluripotent

stem cells: ROCKing between metastable cell states. *Trends Cell Biol,* 21**,** 274-82.

OLSEN, M. W. 1942. Maturation, fertilization, and early cleavage in the hen's egg. J Morphol, 70, 513-533.

ORKIN, S. H., WANG, J., KIM, J., CHU, J., RAO, S., THEUNISSEN, T. W., SHEN, X. & LEVASSEUR, D. N. 2008. The transcriptional network controlling pluripotency in ES cells. Cold Spring Harb Symp Quant Biol, 73, 195-202.

PAIN, B., CLARK, M. E., SHEN, M., NAKAZAWA, H., SAKURAI, M., SAMARUT, J. & ETCHES, R. J. 1996. Long-term in vitro culture and characterisation of avian embryonic stem cells with multiple morphogenetic potentialities. Development, 122, 2339-48.

PANDEY, P., FARBER, R., NAKAZAWA, A., KUMAR, S., BHARTI, A., NALIN, C., WEICHSELBAUM, R., KUFE, D. & KHARBANDA, S. 2000. Hsp27 functions as a negative regulator of cytochrome c-dependent activation of procaspase-3. Oncogene, 19, 1975-81.

PARK, H. J., PARK, T. S., KIM, T. M., KIM, J. N., SHIN, S. S., LIM, J. M. & HAN, J. Y. 2006. Establishment of an in vitro culture system for chicken preblastodermal cells. Mol Reprod Dev, 73, 452-61.

PATTERSON, P. H., KOELKEBECK, K. W., ANDERSON, K. E., DARRE, M. J., CAREY, J. B., AHN, D. U., ERNST, R. A., KUNEY, D. R. & JONES, D. R. 2008. Temperature sequence of eggs from oviposition through distribution: production--part 1. Poult Sci, 87, 1182-6.

PAUL, C., MANERO, F., GONIN, S., KRETZ-REMY, C., VIROT, S. & ARRIGO, A. P. 2002. Hsp27 as a negative regulator of cytochrome C release. Mol Cell Biol, 22, 816-34.

POCCIA, D., WOLFF, R., KRAGH, S. & WILLIAMSON, P. 1985. RNA synthesis in male pronuclei of the sea urchin. Biochim Biophys Acta, 824, 349-56.

PODRABSKY, J. E. & CULPEPPER, K. M. 2012. Cell cycle regulation during development and dormancy in embryos of the annual killifish Austrofundulus limnaeus. Cell Cycle, 11, 1697-704.

PREVILLE, X., SALVEMINI, F., GIRAUD, S., CHAUFOUR, S., PAUL, C., STEPIEN, G., URSINI, M. V. & ARRIGO, A. P. 1999. Mammalian small stress proteins protect against oxidative stress through their ability to increase glucose-6-phosphate dehydrogenase activity and by maintaining optimal cellular detoxifying machinery. Exp Cell Res, 247, 61-78.

PRIESS, J. R. 2005. Notch signaling in the C. elegans embryo. WormBook, 1-16.

QIU, Z. & MACRAE, T. H. 2008a. ArHsp21, a developmentally regulated small heat-shock protein synthesized in diapausing embryos of Artemia franciscana. Biochem J, 411, 605-11.

QIU, Z. & MACRAE, T. H. 2008b. ArHsp22, a developmentally regulated small heat shock protein produced in diapause-destined Artemia embryos, is stress inducible in adults. FEBS J, 275, 3556-66.

QIU, Z., TSOI, S. C. & MACRAE, T. H. 2007. Gene expression in diapause-destined embryos of the crustacean, Artemia franciscana. Mech Dev, 124, 856-67.

RENFREE, M. B. & SHAW, G. 2000. Diapause. Annual review of physiology, 62, 353-375.

RENGARAJ, D., LEE, S. I., PARK, T. S., LEE, H. J., KIM, Y. M., SOHN, Y. A., JUNG, M., NOH, S. J., JUNG, H. & HAN, J. Y. 2014. Small non-coding RNA profiling and the role of piRNA pathway genes in the protection of chicken primordial germ cells. BMC Genomics, 15, 757.

RICHTER, K., HASLBECK, M. & BUCHNER, J. 2010. The heat shock response: life on the verge of death. Mol Cell, 40, 253-66.

RINEHART, J. P., LI, A., YOCUM, G. D., ROBICH, R. M., HAYWARD, S. A. & DENLINGER, D. L. 2007. Up-regulation of heat shock proteins is essential for cold survival during insect diapause. Proc Natl Acad Sci U S A, 104, 11130-7.

ROBINSON, M. D. & OSHLACK, A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology, 11, 1.

ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26, 139-140.

ROESER, T., STEIN, S. & KESSEL, M. 1999. Nuclear beta-catenin and the development of bilateral symmetry in normal and LiCl-exposed chick

embryos. Development, 126, 2955-65.

ROESER, T., STEIN, S. & KESSEL, M. 1999. Nuclear beta-catenin and the development of bilateral symmetry in normal and LiCl-exposed chick embryos. *Development,* 126**,** 2955-65.

ROSSELLO, R. A., CHEN, C. C., DAI, R., HOWARD, J. T., HOCHGESCHWENDER, U. & JARVIS, E. D. 2013. Mammalian genes induce partially reprogrammed pluripotent stem cells in non-mammalian vertebrate and invertebrate species. Elife, 2, e00036.

RUF, T. & GEISER, F. 2014. Daily torpor and hibernation in birds and mammals. Biol Rev Camb Philos Soc.

SAMALI, A., NORDGREN, H., ZHIVOTOVSKY, B., PETERSON, E. & ORRENIUS, S. 1999. A comparative study of apoptosis and necrosis in HepG2 cells: oxidant-induced caspase inactivation leads to necrosis. Biochem Biophys Res Commun, 255, 6-11.

SARETZKI, G., ARMSTRONG, L., LEAKE, A., LAKO, M. & VON ZGLINICKI, T. 2004. Stress defense in murine embryonic stem cells is superior to that of various differentiated murine cells. Stem Cells, 22, 962-71.

SARETZKI, G., WALTER, T., ATKINSON, S., PASSOS, J. F., BARETH, B., KEITH, W. N., STEWART, R., HOARE, S., STOJKOVIC, M., ARMSTRONG, L., VON ZGLINICKI, T. & LAKO, M. 2008. Downregulation of multiple stress defense mechanisms during differentiation of human embryonic stem cells. Stem Cells, 26, 455-64.

SHENG, G. 2014. Day-1 chick development. Dev Dyn, 243, 357-67.

SHI, S. & STANLEY, P. 2006. Evolutionary origins of Notch signaling in early development. *Cell Cycle,* 5**,** 274-8.

SHIN, H., BANG, S., KIM, J., JUN, J. H., SONG, H. & LIM, H. J. 2017. The formation of multivesicular bodies in activated blastocysts is influenced by autophagy and FGF signaling in mice. Sci Rep, 7, 41986.

SHIN, M., ALEV, C., WU, Y., NAGAI, H. & SHENG, G. 2011. Activin/TGF-beta signaling regulates Nanog expression in the epiblast during gastrulation. *Mech Dev,* 128**,** 268-78.

SKROMNE, I. & STERN, C. D. 2001. Interactions between Wnt and Vg1 signalling pathways initiate primitive streak formation in the chick embryo. Development, 128, 2915-27.

SNOOK, R. R., HOSKEN, D. J. & KARR, T. L. 2011. The biology and evolution of polyspermy: insights from cellular and functional studies of sperm and centrosomal behavior in the fertilized egg. Reproduction, 142, 779-92.

SONNA, L. A., FUJITA, J., GAFFIN, S. L. & LILLY, C. M. 2002. Invited review: Effects of heat and cold stress on mammalian gene expression. J Appl Physiol (1985), 92, 1725-42.

STERN, C. D. 2005. The chick; a great model system becomes even greater. Dev Cell, 8, 9-17.

STETLER, R. A., GAO, Y., SIGNORE, A. P., CAO, G. & CHEN, J. 2009. HSP27: mechanisms of cellular protection against neuronal injury. *Curr Mol Med,* 9**,** 863-72.

STREIT, A., BERLINER, A. J., PAPANAYOTOU, C., SIRULNIK, A. & STERN, C. D. 2000. Initiation of neural induction by FGF signalling before gastrulation. Nature, 406, 74-8.

SVOBODA, P. & FLEMR, M. 2010. The role of miRNAs and endogenous siRNAs in maternal-to-zygotic reprogramming and the establishment of pluripotency. EMBO Rep, 11, 590-7.

TADROS, W. & LIPSHITZ, H. D. 2009. The maternal-to-zygotic transition: a play in two acts. Development, 136, 3033-42.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell, 126, 663-76.

TAN, M. H., AU, K. F., YABLONOVITCH, A. L., WILLS, A. E., CHUANG, J., BAKER, J. C., WONG, W. H. & LI, J. B. 2013. RNA sequencing reveals a diverse and dynamic repertoire of the Xenopus tropicalis transcriptome over development. Genome Res, 23, 201-16.

TANG, F., KANEDA, M., O'CARROLL, D., HAJKOVA, P., BARTON, S. C., SUN, Y. A., LEE, C., TARAKHOVSKY, A., LAO, K. & SURANI, M. A. 2007. Maternal microRNAs are essential for mouse zygotic development. Genes Dev, 21, 644-8.

TAO, Q., YOKOTA, C., PUCK, H., KOFRON, M., BIRSOY, B., YAN, D., ASASHIMA, M., WYLIE, C. C., LIN, X. & HEASMAN, J. 2005. Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in Xenopus embryos. Cell, 120, 857-71.

TORLOPP, A., KHAN, M. A., OLIVEIRA, N. M., LEKK, I., SOTO-JIMENEZ, L. M., SOSINSKY, A. & STERN, C. D. 2014. The transcription factor Pitx2 positions the embryonic axis and regulates twinning. Elife, 3, e03743.

TOXOPEUS, J., WARNER, A. H. & MACRAE, T. H. 2014. Group 1 LEA proteins contribute to the desiccation and freeze tolerance of Artemia franciscana embryos during diapause. Cell Stress Chaperones.

TRAPNELL, C., PACHTER, L. & SALZBERG, S. L. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics, 25, 1105-1111.

TSE, Y. C., WERNER, M., LONGHINI, K. M., LABBE, J. C., GOLDSTEIN, B. & GLOTZER, M. 2012. RhoA activation during polarization and cytokinesis of the early Caenorhabditis elegans embryo is differentially dependent on NOP-1 and CYK-4. Mol Biol Cell, 23, 4020-31.

TSUNEKAWA, N., NAITO, M., SAKAI, Y., NISHIDA, T. & NOCE, T. 2000. Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells. Development, 127, 2741-50.

UNTERGASSER, A., CUTCUTACHE, I., KORESSAAR, T., YE, J., FAIRCLOTH, B. C., REMM, M. & ROZEN, S. G. 2012. Primer3--new capabilities and interfaces. Nucleic Acids Res, 40, e115.

VAN DE LAVOIR, M. C., MATHER-LOVE, C., LEIGHTON, P., DIAMOND, J. H., HEYER, B. S., ROBERTS, R., ZHU, L., WINTERS-DIGIACINTO, P., KERCHNER, A., GESSARO, T., SWANBERG, S., DELANY, M. E. & ETCHES, R. J. 2006. High-grade transgenic somatic chimeras from chicken embryonic stem cells. Mech Dev, 123, 31-41.

VILLENEUVE, T. S., MA, X., SUN, Y., OULTON, M. M., OLIVER, A. E. & MACRAE, T. H. 2006. Inhibition of apoptosis by p26: implications for small heat shock protein function during Artemia development. Cell Stress Chaperones, 11, 71-80.

WANG, J., FAN, H. C., BEHR, B. & QUAKE, S. R. 2012. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. Cell, 150, 402-12.

WANG, S. H., CHENG, C. Y., TANG, P. C., CHEN, C. F., CHEN, H. H., LEE, Y. P. & HUANG, S. Y. 2013. Differential gene expressions in testes of L2 strain Taiwan country chicken in response to acute heat stress. Theriogenology, 79, 374-82 e1-7.

WANG, S. H., CHENG, C. Y., TANG, P. C., CHEN, C. F., CHEN, H. H., LEE, Y. P. & HUANG, S. Y. 2015. Acute heat stress induces differential gene expressions in the testes of a broiler-type strain of Taiwan country chickens. PLoS One, 10, e0125816.

WELTY, J. C. 1982. The Life of Birds, New york : Saunders College Publishing.

WIEKOWSKI, M., MIRANDA, M. & DEPAMPHILIS, M. L. 1991.

Regulation of gene expression in preimplantation mouse embryos: effects of the zygotic clock and the first mitosis on promoter and enhancer activities. Dev Biol, 147, 403-14.

WINKLER, D. W. & WALTERS, J. R. 1983. The determination of clutch size in precocial birds. Current ornithology. Springer.

WU, J., HUANG, B., CHEN, H., YIN, Q., LIU, Y., XIANG, Y., ZHANG, B., LIU, B., WANG, Q., XIA, W., LI, W., LI, Y., MA, J., PENG, X., ZHENG, H., MING, J., ZHANG, W., ZHANG, J., TIAN, G., XU, F., CHANG, Z., NA, J., YANG, X. & XIE, W. 2016. The landscape of accessible chromatin in mammalian preimplantation embryos. Nature, 534, 652-7.

XUE, Z., HUANG, K., CAI, C., CAI, L., JIANG, C. Y., FENG, Y., LIU, Z., ZENG, Q., CHENG, L., SUN, Y. E., LIU, J. Y., HORVATH, S. & FAN, G. 2013. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature, 500, 593-7.

YAN, L., YANG, M., GUO, H., YANG, L., WU, J., LI, R., LIU, P., LIAN, Y., ZHENG, X., YAN, J., HUANG, J., LI, M., WU, X., WEN, L., LAO, K., LI, R., QIAO, J. & TANG, F. 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol, 20, 1131-9.

YANG, J., RYAN, D. J., WANG, W., TSANG, J. C., LAN, G., MASAKI, H., GAO, X., ANTUNES, L., YU, Y., ZHU, Z., WANG, J., KOLODZIEJCZYK, A. A., CAMPOS, L. S., WANG, C., YANG, F., ZHONG, Z., FU, B., ECKERSLEY-MASLIN, M. A., WOODS, M., TANAKA, Y., CHEN, X., WILKINSON, A. C., BUSSELL, J., WHITE, J., RAMIREZ-SOLIS, R.,

REIK, W., GOTTGENS, B., TEICHMANN, S. A., TAM, P. P. L., NAKAUCHI, H., ZOU, X., LU, L. & LIU, P. 2017a. Establishment of mouse expanded potential stem cells. Nature, 550, 393-397.

YANG, Y., LIU, B., XU, J., WANG, J., WU, J., SHI, C., XU, Y., DONG, J., WANG, C., LAI, W., ZHU, J., XIONG, L., ZHU, D., LI, X., YANG, W., YAMAUCHI, T., SUGAWARA, A., LI, Z., SUN, F., LI, X., LI, C., HE, A., DU, Y., WANG, T., ZHAO, C., LI, H., CHI, X., ZHANG, H., LIU, Y., LI, C., DUO, S., YIN, M., SHEN, H., BELMONTE, J. C. & DENG, H. 2017b. Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. Cell, 169, 243-257 e25.

YARTSEVA, V. & GIRALDEZ, A. J. 2015. The Maternal-to-Zygotic Transition During Vertebrate Development: A Model for Reprogramming. Curr Top Dev Biol, 113, 191-232.

YEO, J. C. & NG, H. H. 2013. The transcriptional regulation of pluripotency. Cell Res, 23, 20-32.

YU, J., VODYANIK, M. A., SMUGA-OTTO, K., ANTOSIEWICZ-BOURGET, J., FRANE, J. L., TIAN, S., NIE, J., JONSDOTTIR, G. A., RUOTTI, V., STEWART, R., SLUKVIN, II & THOMSON, J. A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science, 318, 1917-20.

ZAMIR, E., KAM, Z. & YARDEN, A. 1997. Transcription-dependent induction of G1 phase during the zebra fish midblastula transition. Mol Cell Biol, 17, 529-36.

ZENG, F. & SCHULTZ, R. M. 2005. RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. Dev Biol, 283, 40-57.

ZERNICKA-GOETZ, M., MORRIS, S. A. & BRUCE, A. W. 2009. Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. Nat Rev Genet, 10, 467-77.

ZHANG, B., ZHENG, H., HUANG, B., LI, W., XIANG, Y., PENG, X., MING, J., WU, X., ZHANG, Y., XU, Q., LIU, W., KOU, X., ZHAO, Y., HE, W., LI, C., CHEN, B., LI, Y., WANG, Q., MA, J., YIN, Q., KEE, K., MENG, A., GAO, S., XU, F., NA, J. & XIE, W. 2016. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature, 537, 553-557.

ZHANG, G., LI, C., LI, Q., LI, B., LARKIN, D. M., LEE, C., STORZ, J. F., ANTUNES, A., GREENWOLD, M. J., MEREDITH, R. W., ODEEN, A., CUI, J., ZHOU, Q., XU, L., PAN, H., WANG, Z., JIN, L., ZHANG, P., HU, H., YANG, W., HU, J., XIAO, J., YANG, Z., LIU, Y., XIE, Q., YU, H., LIAN, J., WEN, P., ZHANG, F., LI, H., ZENG, Y., XIONG, Z., LIU, S., ZHOU, L., HUANG, Z., AN, N., WANG, J., ZHENG, Q., XIONG, Y., WANG, G., WANG, B., WANG, J., FAN, Y., DA FONSECA, R. R., ALFARO-NUNEZ, A., SCHUBERT, M., ORLANDO, L., MOURIER, T., HOWARD, J. T., GANAPATHY, G., PFENNING, A., WHITNEY, O., RIVAS, M. V., HARA, E., SMITH, J., FARRE, M., NARAYAN, J., SLAVOV, G., ROMANOV, M. N., BORGES, R., MACHADO, J. P., KHAN, I., SPRINGER, M. S., GATESY, J., HOFFMANN, F. G., OPAZO, J. C., HASTAD, O., SAWYER, R. H., KIM, H., KIM, K. W., KIM, H. J., CHO, S., LI, N., HUANG, Y., BRUFORD, M. W., ZHAN, X., DIXON, A., BERTELSEN, M. F., DERRYBERRY, E., WARREN, W., WILSON, R. K., LI, S., RAY, D. A.,

GREEN, R. E., O'BRIEN, S. J., GRIFFIN, D., JOHNSON, W. E., HAUSSLER, D., RYDER, O. A., WILLERSLEV, E., GRAVES, G. R., ALSTROM, P., FJELDSA, J., MINDELL, D. P., EDWARDS, S. V., BRAUN, E. L., RAHBEK, C., BURT, D. W., HOUDE, P., ZHANG, Y., et al. 2014a. Comparative genomics reveals insights into avian genome evolution and adaptation. Science, 346, 1311-20.

ZHANG, G., RAHBEK, C., GRAVES, G. R., LEI, F., JARVIS, E. D. & GILBERT, M. T. 2015. Genomics: Bird sequencing project takes off. Nature, 522, 34.

ZHANG, H. M., CHEN, H., LIU, W., LIU, H., GONG, J., WANG, H. & GUO, A. Y. 2012. AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res, 40, D144-9.

ZHANG, W. L., HUITOREL, P., GLASS, R., FERNANDEZ-SERRA, M., ARNONE, M. I., CHIRI, S., PICARD, A. & CIAPA, B. 2005. A MAPK pathway is involved in the control of mitosis after fertilization of the sea urchin egg. Dev Biol, 282, 192-206.

ZHANG, Y., DUAN, X., CAO, R., LIU, H. L., CUI, X. S., KIM, N. H., RUI, R. & SUN, S. C. 2014b. Small GTPase RhoA regulates cytoskeleton dynamics during porcine oocyte maturation and early embryo development. Cell Cycle, 13, 3390-403.

# SUMMARY IN KOREAN

조류는 여러 연구, 특히 발생학을 위한 가치 있는 모델 시스템으로 활용되어왔다. 또한, 척추 동물의 비교유전체학 분야에서 중요한 위치이자 많은 연구 분야에서의 관련성으로 인해 Bird10K 프로젝트 (B10K) 가 2014~2015년에 시작되었다. 이 프로젝트의 일환으로, 조류의 계통 발생 계급과 비행 및 기능 적응을 위한 비교유전체학도 분석되었다. 역사적으로 의미 있고 이론적으로 우수하지만, 방란 전 배아의 접근이 어려워 조류 배아 발생의 초기 및 중요 이벤트를 조사하는 지놈 전체적인 전사 발현에 관한 연구는 아직 수행되지 못했다. 이러한 측면에서 본 연구에서는 닭의 난자, 접합자와 Eyal-giladi 및 Kochav stage I (EGK.I) 에서 EGK.X 까지의 자궁 내 배아를 포함한 방란 전 배아를 비 침습적 방법을 이용하여 획득하고, 최초 전사체 데이터를 제시했다. 또한, 본 연구의 다중 오믹 접근법은 닭의 수정 시 첫 번째 전사 활성화를 조사했다. 또한, 조류 특이적 작은 열쇼크단백질 (sHSP) 인 HSP25 를 밝혀내고, 방란 후 일시적 발달 중단 상태인 조류 포배 휴면기에서의 역할을 발견했다.

첫 번째 연구는 초기 닭 배아의 전체 전사체적 정보를 생성하고 분석하기 위해 수행되었다. 이번 연구에서는 닭의 난소 및 난관으로부터 방란 전 배아 총 137 개 수집하고 RNA 시퀀싱 (RNA-seq)을 수행했다. 이러한 결과로부터 Notch, MAPK, Wnt 및 TGF-β 신호 전달이 두 차례의 zygotic genome activation (ZGA)에 의해 발현되어 cleavage 및 area pellucida formation 시기를 별개의 프로그램으로 각각 조절하였다. 또한, EGK stage 를 사람과 쥐의 초기 발달 시기와 비교하여 닭 특유의 신호전달 체계의 특징을 강조하였고, 그들의 발현 유사도가 ZGA 전후로 점차 비슷해짐을 발견하였다. 이러한 발견을 통해 조류의 배 발생과 양막류 간의 비교를 지놈 전체적인 발현 측면에서 제공한다.

생산된 전사체 데이터를 바탕으로, 척추동물에서 배아 정체성을 확립하는 중요한 과정인 maternal-to-zygotic transition (MZT)을 확인했다. semi-supervised clustering에 따른 RNA-seq 데이터에 대한 동시 발현 분석에 기초하여, 두 차례의 ZGA 를 매개로 한 MZT 가 자궁 내 발달에서 관찰되었고, 전사 및 번역 관련 유전자 발현의 역동성이 발견되었다. 또한 기능분석을 통해, cleavage 와 area pellucida formation 시기의 발달 프로그램의 전환이 뚜렷하게 발견되었다. 마지막으로 후생유전학적 리프로그래밍과 miR-302s 의 발현은 초기 닭 발달 과정에서 EGK.VIII 가 명확한 MZT가 진행되는 시기임을 제안하였다. 이러한 연구는 MZT 조절 관점에서 척추 동물 간의 진화론적 연결을 제공할 것으로 기대한다.

포유류의 유도 다능성 줄기세포 (iPSCs)와 확장된 다능성 줄기세포 (EPSCs)는 난자와 초기 배아의 전사 인자의 연구를 토대로 생산되었다. 그러나 현재까지도 조류의 초기 발달 과정에서의 인자들을 조사하지 않은 채, 포유류의 전사인자만을 이용하여 iPSC 유사세포로 완전히 리프로그래밍되지 못했다. 세 번째 연구에서는 닭 초기 발달에서 전사체 수준에서의 유전자 발현의 전환과 관련 전사인자를 특성화하였다. 결론적으로, 두 차례의 ZGA에 의해 발현되는 전사인자들이 각각 cleavage와 area pellucida formation 에 관련되어 있었다. 이 결과로부터, 리프로그래밍이 가능한 다수의 조류 특이적 전사인자를 발굴하였고, 포유류의 전능성과 비슷하고 더 높은 수준의 잠재성을 지니는 cleavage 시기를 조류 리프로그래밍에서 지향해야 함을 주장한다.

조류 종에서 ZGA 를 더 깊이 연구하기 위해 우리는 닭에서 ZGA 개시에 대하여 상세히 조사했다. 전사체적 접근과 인트론 지역 발현을 통해 두 차례의 ZGA는 수정 이후와 EGK.V 에서 일어나는 것을 확인하였다. 또한, 다중 오믹 접근을 통해 품종 특이적 단일 뉴클레오티드 다형성 (SNPs)을 토대로 한 대립유전자 발현으로 첫 ZGA 의 발현을 추적하였다.

놀랍게도 닭의 접합자 시기에서는 부계의 전핵이나 다정자의 핵들이 존재함에도 불구하고 모체 유전체의 활성만이 나타났다. 이러한 모체 유전자는 cleavage 시기를 조절하고, MZT 를 거치며 두 대립 유전자의 발현으로 대체되었다. 이 결과를 통해 조류의 접합자가 모체 유전자만 발현하도록 하는 메커니즘이 존재하고 이는 부계 유전자의 발현으로 인한 불균형적인 지놈 기여 또는 유전적 불안정성을 방어로 인한 것으로 주장하였다.

마지막으로, 스트레스 저항성을 지니는 조류 포배의 특이적 sHSP 를 밝혀냈다. 닭 HSP25 는 닭 포배에서 특이적으로 발현하며 저온 저장에 의해 발현이 증가하였다. 아미노산 서열 비교와 메추리, 핀치새의 발현을 토대로 HSP25 는 조류에서 보존되어 있는 유전자임을 밝혔다. HSP25 의 발현저해를 통해, 전능성 유전자의 발현이 감소하였다. 또한 HSP25 는 포배 세포에서 항 세포사멸, 항 산화, 친 오토파지 역할을 돕고 있었다. 종합적으로, 조류 HSP25 는 조류 포배 세포를 보호하고 환경 스트레스에 빠르게 대응하는 데 중요한 역할을 한다.